



North Coast and Cascades Monitoring Network of the National Park Service

Temporal Sampling Frames: Summary of a Workshop

Andrea Woodward

USGS Forest & Rangeland Ecosystem Science Center Olympic Field Station, 600 E. Park Avenue, Port Angeles, WA 98236

Note: This report summarizes the results of a workshop held in Port Angeles, WA, and is meant to describe ideas presented at the workshop. Workshop attendees have been given the opportunity to review the concepts and interpretations presented, but the contents have not been subjected to outside peer-review because it was not deemed appropriate. Without outside peer-review, however, the contents cannot be approved by the USGS director. They are provided with the understanding that **anyone interested in citing this information should contact the authors before using** and should carefully describe the conclusions as those of the workshop attendees. Conclusions drawn from, or actions undertaken on the basis of, this information are the sole responsibility of the user.

U.S. Department of the Interior
U.S. Geological Survey

Introduction

The USGS Forest and Rangeland Ecosystem Science Center Olympic Field Station has been working with the North Coast and Cascades Network (NCCN) of National Parks to develop a long-term ecological monitoring program. One important role of monitoring is to provide managers with a scientific basis for their decisions. As park management decisions come under closer and closer scrutiny by the public, park managers must have a solid justification for their decisions. If the results of monitoring determine a need for management action, those results must stand up to scientific peer-review. Among the elements that will contribute to scientific credibility is a statistically valid sampling scheme that includes a large number of spatially and temporally dispersed samples.

The NCCN consists of seven parks in northwestern Oregon and western Washington. They include four lowland, cultural and historic parks (i.e., Ebey's Landing National Historical Reserve, Fort Clatsop National Memorial, Fort Vancouver National Historic Site, San Juan Island National Historic Park) and three mountainous natural area parks (i.e., Olympic and Mount Rainier National Parks, and North Cascades National Park Complex). Developing sampling designs, both the spatial and temporal components, is most challenging at the large parks because of their sheer size, access difficulties in wilderness areas with steep terrain, and diversity of resources.

Among the challenges for monitoring these national parks is the need for temporal designs that will maximize the ability to describe status and trend at the same time. To address these issues, USGS convened a workshop with the following goals:

1. Review survey design principles
2. Summarize temporal monitoring designs and their terminology
3. Consider practical considerations in designing panel surveys, and
4. Provide practical recommendations on survey designs for selected monitoring programs in NCCN.

The workshop was attended by several invited biometricians with experience working with NPS and other federal agencies to develop long-term monitoring programs. Monitoring specialists from USGS and NPS who work with NCCN, as well as resource management staff members from the three large parks also attended. The workshop began with presentations by three invitees covering the first three objectives. These were followed by descriptions of three selected monitoring projects presented by NCCN and USGS staff members. The monitoring projects were chosen to represent a range of monitoring challenges. The group discussed the sample plan for each project to illustrate the application of general sampling principles, not necessarily to develop detailed sampling designs.

Overview of Survey Designs (presented by Tony Olsen, EPA, Corvallis OR)

Survey designs are necessary tools for effective inventory and monitoring, such as the program under development by the NPS to assess the status and trends of natural resources. *Inventories* are defined as an extensive “point in time” effort to determine the location, condition, or itemized list of resources. The specific goals of the NPS inventory are to describe the distribution and abundance of certain species, and they exemplify the problems of stating goals imprecisely. The weaknesses of this statement are that the parks do not have the resources to visit all sites, yet “distribution” implies a map and requires visits to all sites. Also, the requirement for “abundance” estimates does not indicate whether absolute or relative abundance or presence/absence is required.

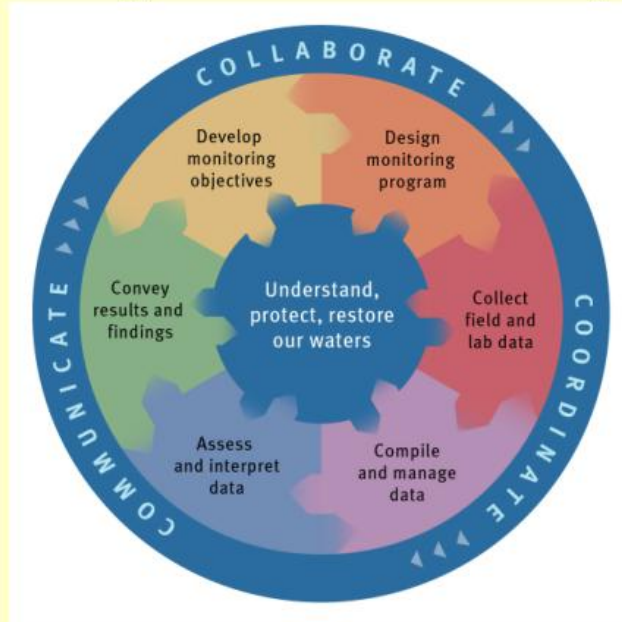
Monitoring is the collection of repeated observations or measurements to evaluate changes in condition, or progress toward meeting a management objective (Elzinga et al. 1998).

Monitoring programs are prone to several weaknesses:

- Monitoring results are not directly tied to management decision-making – there is need to get information to managers and other key audiences in a timely fashion.
- Objectives for monitoring are not clear, precisely stated and understood – the EMAP and NAWQA programs both had the overall goal of monitoring status and trends in water quality, but their specific objectives and questions were different. Describing status and trends is much too vague an objective to address sampling design issues.
- Measurement protocols, survey designs, and statistical analyses for monitoring become scientifically out of date and eventually obsolete (i.e., every 10-15 years). You need to plan for changes in technology and statistical techniques and be sure to allow time to develop a calibration curve to compare the existing protocol with proposed changes.

These pitfalls can be avoided by seeing monitoring as an information system, where the pieces of the monitoring framework are designed and implemented to fit together so they can become the central organizing approach to managing natural resources (Figure 1). This presentation will focus on the monitoring design and data collection.

Monitoring is an Information System



28

Figure 1. Monitoring as an information system (from Tony Olsen's presentation). Mentally replace 'waters' in the center piece with 'resources' for purposes of the NPS.

Briefly, developing a survey design for monitoring involves the following steps:

- Clearly state quantitative objectives
- Explicitly define the target population
- Construct the sample frame to represent the target population
- Decide on the survey design
- Select sites using the survey design
- Use a statistical analysis that matches the survey design.

We will cover each step in detail.

Monitoring Objective

Non-specific monitoring objectives are a weakness of most monitoring programs. Often, objectives are not clear, precisely stated, understood, or linked to management decisions and reporting requirements. Objectives determine the monitoring design, and it is common to have multiple objectives. Importantly, the objectives should be prioritized because objectives compete for samples. Two examples of moving from a generally stated monitoring question to specific, quantitative objectives are:

Water Quality Example:

- What is the overall quality of waters in the park?

- What is the overall quality of streams with flowing water during summer in the park?
- What is the biological quality of streams with flowing water during summer in the park?
- How many km of streams with flowing water during summer are impaired, non-impaired, or marginally impaired within the park? (where 'summer,' 'stream with flowing water,' and 'impairment' are quantitatively defined; are perennial streams included? How are they defined?)

Amphibian Example:

- What are the distribution and abundance of the northwestern salamander (*Ambystoma gracile*: AMGR) in Olympic National Park and how are these changing over time?
- How many ponds in ONP have one or more AMGRs present?
- What percent of ponds in ONP have AMGRs present and what is the trend trajectory in percent of ponds occupied? Or,
- What is the probability of AMGR occupancy for all ponds in ONP?

Defining Elements of the Target Population

The target population is the resource for which you want information, and should follow from the monitoring objective. The definition must be clear and understandable to users, especially the field crew that must be able to determine if a particular site should be included in the sample. It is usually more difficult than expected to define the elements that make up the target population. For example, if the target population is ponds, with an individual pond as the population element, you would want to measure something about the entire pond. However the definition of this target population is not clear unless you also define the difference between a pond and a wetland, and at what size a pond becomes a lake, because wetlands and lakes are likely to be sampled differently than ponds. For example, lakes may require multiple samples (e.g., one sample is probably not adequate to describe Lake Chelan; you might want to sample it according to urban, agricultural and pristine influences). Analogously for streams, are the population elements reaches or points along the stream? How do you define a wadeable stream? Do you want to sample Hydrologic Unit Codes (HUCs; hierarchical classification of sub-watersheds mapped by the Forest Service)? How do you handle HUCs that cross park boundaries? One helpful tool for answering these questions is to realize that the elements of the target population are closely related to the unit of replication.

Subpopulations also deserve careful consideration because they have a significant impact on the monitoring design. Subpopulations (also called 'domains') are critical areas or elements having high priority for monitoring. They must be defined early in the design process to insure the design provides adequate sample sizes, especially if the domain is a small part of the total population. This can be achieved by explicit stratification, or implicitly when other requirements of the sample design automatically provide an adequate sample (e.g., a design giving higher probability of selection to high elevation areas would automatically increase the sample of subalpine ecosystems). It may also be true that some subpopulations cannot be defined in advance and therefore cannot be accounted for in the survey design. In this case, domain members can be identified after data are collected, but the sample may be insufficient.

Sample Frame

The sample frame is a representation of the target population that is used to select sample sites. It consists of all sample units that are potential members of the sample. For example, the sample frame may include all stream reaches for stream sampling, or a grid of potential sample sites for vegetation.

Unfortunately, the sample frame is almost never an exact representation of the target population. Due primarily to limitations of mapping tools, some elements are usually under-covered while others are over-covered (Figure 2). For example, one might look at a GIS coverage that includes all streams as a sample frame. In it, perennial streams may not be well mapped (under-coverage) and some features that are not streams are included (over-coverage). Consequently a survey design based on this frame will provide limited information for perennial streams, and other information is needed to define the perennial domain. When a sample is drawn from the sample frame some points are unavailable due to technical limitations, physical barriers, limited access to private property, etc. Consequently, some points are unavailable or inappropriate to sample, thereby limiting the actual sampled population to the subset of the sample frame that is not subject to those limitations.

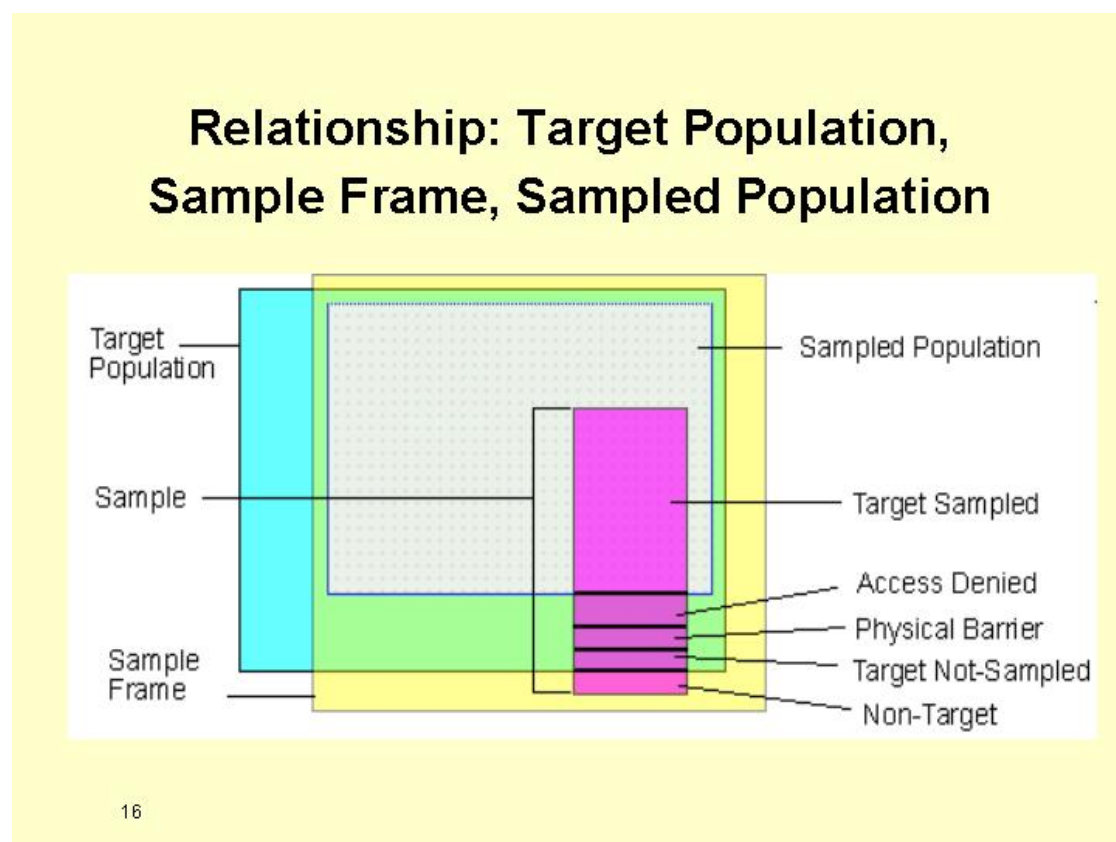


Figure 2. Correspondence among target population, sample frame, and the sample (from Tony Olsen's presentation). Target population = goal; Sample frame = best definition of target using imperfect mapping tools; Sample = elements chosen from sample frame; Target sampled = the part of the sample that is accessible and is actually part of the target population; Sampled population = the part of the target population represented by the target sampled.

Survey Design

Choosing a 'survey' design from among all possible statistical designs as the appropriate statistical design for monitoring programs comes from the following considerations:

- The goal of monitoring is to obtain a representative sample from which to infer changes to the entire target population. "Representative" or "authoritative" sites, which are subjectively chosen based on expert judgment, provide no valid inference. In this case, any inference depends on the judgment of the expert and is not replicable and consequently unscientific. Valid inference requires a statistical design.
- There are three types of statistical designs, but only survey designs have appropriate properties for monitoring:
 - Experimental designs have the objective of understanding ecological processes by testing the effectiveness of various treatments that are applied by the experimenter who randomly allocates treatments to units. Experimental designs do not provide estimates for populations or information on the current state of the ecosystem.
 - Observational studies have the objective of understanding ecological processes, but the experimenter cannot control or apply treatments. Instead, the designs usually involve factor space designs (e.g., gradient studies) where sites are located to take advantage of a 'natural' experiment. Researchers must take advantage of a limited number of available sites.
 - Survey designs have the objective of making estimates or inference for some specified finite population, such as a national park. The researcher uses either a census (sample every element of the target population; not often feasible) or a probability survey (sample a subset of the target population in a probabilistic way that allows inference to the entire target population). The ability to make inferences about ecological processes is limited because the effects of other factors cannot be controlled by randomly assigning treatments to experimental units.

Note: A sample design has both spatial and temporal components. Survey designs relate to the process of selecting sites in space, while temporal designs describe how samples are distributed through time. Response designs address the particular methods applied at a site to measure a response. The response design may have a temporal component that defines when the sample must be taken (i.e., index period, such as season). While this workshop focuses specifically on the temporal distribution of sample units, this cannot be done without considering the associated spatial sampling design.

There are three categories of spatial survey designs:

- Simple random sample – does not result in an evenly distributed sample because random samples are often clumped.
- Systematic sample – either using a regular grid or regular spacing on a linear resource (e.g., streams). It provides domain elements in the proportion they naturally occur thereby over-sampling the common elements and under-sampling the rare ones.
- Spatially balanced design – a combination of the simple random and the systematic sample to guarantee that all possible samples are geographically evenly dispersed across the target population (e.g., Generalized Random Tessellation Stratified [GRTS]; Steven 1997).

Unfortunately, these basic designs are not sufficient to provide adequate samples of rare subpopulations (domains) or to address administrative restrictions, operational costs, and resources restricted to a specific rare habitat. Consequently, stratified designs are often adopted, which divide the population of interest into a number of strata such that each unit belongs to one and only one stratum. Separate samples are selected in each stratum. Strata are used for several reasons: administrative or operational convenience (e.g., when states need to be operationally independent), to provide different designs appropriate for different portions of the target populations (e.g., design for extensive wetlands like the Everglades may be different than for prairie pothole wetlands), to maximize the precision for a fixed budget by considering different costs of access, or to increase precision by constructing homogeneous strata. The main drawback of stratification is that people often want to change the strata at a later time, and this cannot be legitimately accomplished.

Other more complex survey designs exist:

- Spatial strata random sample – an alternative way to spatially balance the sample (e.g., randomly sample within strata defined by elevation)
- Unequal probability sample – an alternative to stratification that requires auxiliary information (e.g., assign probability of selection for an element based on its distance from a trail)
- Cluster sample – sample several sites in clusters. This can decrease the cost of field operations; however, the independent sample size is only the number of clusters rather than the total number of plots.
- Multiple stage sample – a way to decrease the cost of sample frame construction (e.g., USFS wanted to sample 6th field HUCs (subsets of 5th field HUCs) at a time when only 5th field HUCs had been mapped. Their solution was to sample 5th field HUCs first, then randomly sampled 6th field HUCs within the selected 5th field HUCs after mapping was complete).

Tony Olsen recommends using an unequal probability sample over a stratified sample. One important advantage is that unequal probability samples provide greater flexibility into the future because each sample has a known probability of selection that is not tied to the original boundaries among strata, but this approach does increase the complexity of the analyses somewhat.

Site Selection

Sites for monitoring are selected according to the survey design adopted. Measurements taken from the sites must be analyzed according to the survey design.

Statistical Analysis

The specific objective will determine the survey design, which will, in turn, determine the appropriate statistical analysis. For example, the program PRESENCE was developed to calculate proportion of sites occupied, and is applicable for amphibians and birds. This software assumes simple random sampling – if an unbalanced survey design is used, the software routines must be modified.

Overview of Panel Designs (presented by Trend McDonald, WEST, Inc. based on McDonald 2003)

National Park Service monitoring goals include the understanding of both status and trends for park resources. These are difficult (expensive) to achieve concurrently because status requires spatially distributed samples and trend requires frequent visits. In general, ‘panel’ sampling designs are used to allocate sampling both spatially and temporally to effectively manage the trade-off between status information and trend information. To describe types of panel designs and their attributes, this presentation will answer three questions: 1) what is trend, 2) what is a panel design, and 3) what types of panel designs are found in the literature? As we attempt to design monitoring for the NPS, it will be useful to have an introduction to panel designs and terminology.

What is trend?

There are two types of trend, individual trend and net trend, and populations experience both at once. Individual trend refers to a consistent pattern (direction) of change in a parameter describing an individual member of a population. Examples include:

- A timber company may be interested in salmon occurring in a stream that is scheduled to be cut over, rather than salmon across all ownerships.
- All vegetation dies on a plot between time 1 and time 2.
- Fish move into and populate a certain stream reach.
- An individual elk contracts a disease between time 1 and time 2.

Net trend is a change in a parameter that summarizes the status of all members of a population. Examples include:

- A net change in vegetation cover would be change in the average cover across a number of plots.
- Change in the ratio of infected elk to healthy elk in a population.

A population can experience net change without individual change (e.g., emigration or immigration may result in individuals with particular attributes coming or going from the population, but there may be no change in the individuals originally in the population). Likewise, a population can experience individual change without net change (e.g., if some members increase and others decrease their response for a parameter, they may cancel in the net estimate). Populations can experience individual trend and net trend at the same time. Whether or not a change is individual or net depends on the definition of the population and the definition of the measured parameters. Most monitoring programs are interested in net trend; therefore our discussion will focus on it.

Panel Designs

A **panel** consists of a group of population units that are always all sampled during the same sampling occasion. They are defined spatially by the **membership design**, which is the plan by which population units become members of panels; they are sampled temporally according to the **revisit design**, which is the plan by which panels are sampled in time.

Membership Designs. Using examples from Denali National Park, two membership schemes can be illustrated. The first scheme divides a systematic grid of 81 points (5 km spacing) into groups of 9 sample points organized in 3 x 3 blocks (Figure 3). Some points fall outside of the Park and are not sampled. One point from each block is included in each panel, making each panel a systematic sample of the entire park. One panel is measured each year, so each sample point is re-measured every 9 years. This scheme can be described as 9 ‘interpenetrating’ panels and has the advantage of giving an estimate with inference to the entire park each year. The major disadvantage of this design is the high travel costs necessitated by reaching all parts of the park each year.

Membership Scheme 1

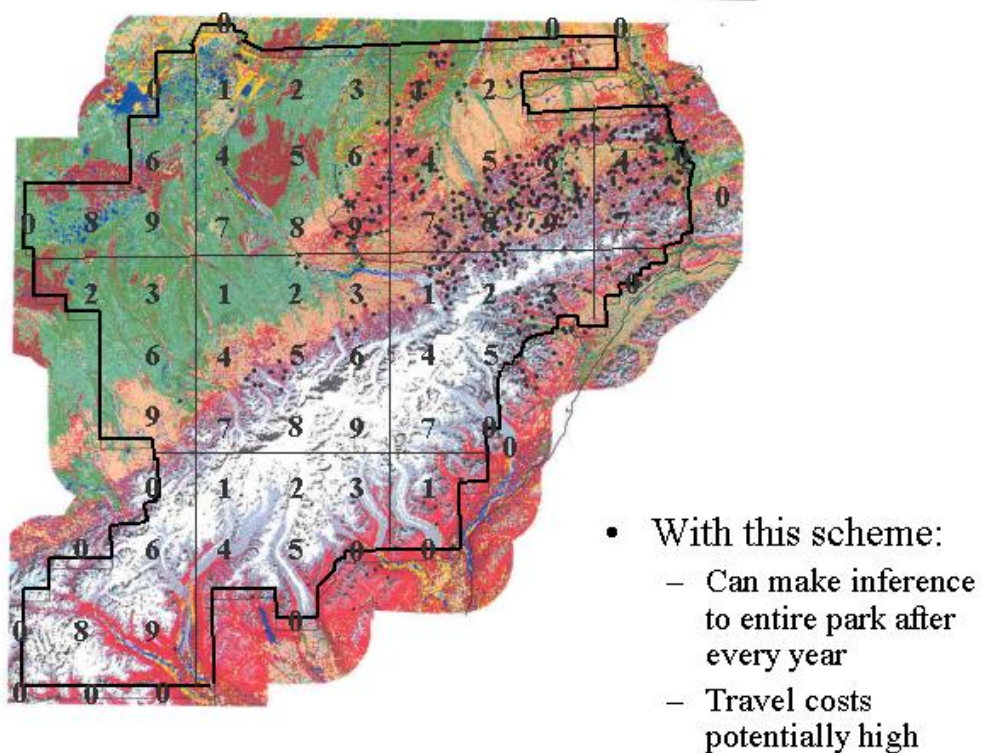
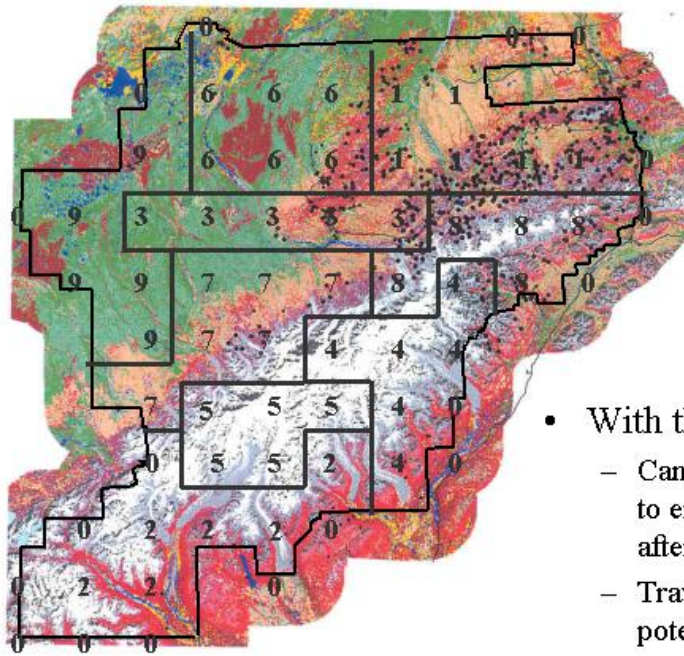


Figure 3. Interpenetrating membership design for sampling points on a systematic grid at Denali National Park. Each panel is a systematic sample from the entire park (all plots with the same number are in the same panel; modified from Trent McDonald’s presentation).

A second membership design would divide the same systematic grid into strata corresponding to high mountains, foothills, flats, etc. (Figure 4). This design is called a stratified systematic sample and has the advantage of lower travel costs, but each year’s estimate can only be inferred to the stratum sampled that year.

Membership Scheme 2



- With this scheme:
 - Can make inference to entire park only after year 9
 - Travel costs potentially lower

Figure 4. Stratified systematic membership design for sampling points on a systematic grid at Denali National Park. Each panel is a stratum. After one full cycle we have a stratified systematic sample (modified from Trent McDonald's presentation).

Revisit Designs. There are many ways the 9 panels in the Denali example could be visited. Revisit Scheme 1 shows each panel visited for one year and not revisited until all of the other panels have been measured (Figure 5). Only years at 5-year intervals can be compared because the design is not 'connected'. A design is statistically connected if it is possible to write a contrast of observations that estimates the difference between any two years. Model assumptions are required to estimate trends from a disconnected design. For example, with Revisit Scheme 1, one could average the 10-year differences from all panels to estimate an average rate of change (assuming a linear model).

Revisit Scheme 1

	Year																	
Panel	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	x									x								
2		x									x							
3			x									x						
4				x									x					
5					x									x				
6						x									x			
7							x									x		
8								x									x	
9									x									x

(5 to 6 plots per year)

Figure 5. Revisit Scheme 1 for panels at Denali NP (from Trent McDonald's presentation).

Revisit Scheme 2 describes each panel being visited for two consecutive years and then rested for seven years (Figure 6). This type of design is 'connected' because there is a statistical linkage (replication) between consecutive years, even though different panels provide linkage for each comparison (i.e., Panel 1 links years 1 & 2, Panel 2 links years 2 & 3, etc.).

'Connection' means that regular linear models can be used to compare years. Measuring each point in successive years allows the direct estimation of difference between successive years. Comparing other years is less precise because more complex contrasts are required. For example, the difference between years 1 and 2 is estimated as $y_{11} - y_{12}$ with variance $2s^2$, where $y_{\text{panel}, \text{year}}$ is an observation and s^2 is the variance of an observation. The variance of a contrast Σa is $\Sigma a^2 s^2$ where a is the coefficient of each observation in the contrast. The difference between years 1 and 5 is estimated as $y_{11} - y_{12} + y_{22} - y_{23} + y_{33} - y_{34} + y_{44} - y_{45}$ with variance $8s^2$ (because the coefficient for each observation is either 1 or -1). This scheme requires twice as much sampling each year as Scheme 1 unless fewer units are included in each panel.

Revisit Scheme 2

	Year																	
Panel	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	x	x								x	x							
2		x	x								x	x						
3			x	x								x	x					
4				x	x								x	x				
5					x	x								x	x			
6						x	x								x	x		
7							x	x								x	x	
8								x	x								x	x
9	x								x	x								x

(10 to 12 plots per year)

Figure 6. Revisit Scheme 2 for panels at Denali NP (from Trent McDonald's presentation).

Revisit Scheme 3 describes Panel 1 being visited every year, and the other 8 panels visited for one year and rested for nine years (Figure 7). This is called a simple split-panel design where groups of panels are visited on different schedules. Sampling some sites every year causes this design to be highly connected and also requires twice the annual effort as Revisit Scheme 1 assuming the same number of sites are included in each panel. Direct comparisons among all 11 years can be made from panel 1, but the comparison is based on only 1 panel. Other panels can only be used to estimate the difference between years the years in which they are sampled. Consequently, Scheme 2 is better (more precise) for estimating difference between successive years, but Scheme 3 panel 1 is better for estimating the multi-year population trajectory. Panels 2-9 are used to estimate status but not trends.

Revisit Scheme 3

	Year																	
Panel	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
2	x								x								x	
3		x								x								x
4			x								x							
5				x								x						
6					x								x					
7						x								x				
8							x								x			
9								x								x		

(10 to 12 plots per year)

Figure 7. Revisit Scheme 3 (simple split-panel design) for panels at Denali National Park.

It is easy to imagine that much more complicated designs than these are possible. A short hand notation has been developed to describe rotating panel designs in less space than required by a revisit table (McDonald 2003). The notation describes the number of panels revisited according to a visit/rest pattern:

- $[x-y]$ = all panels are sampled for x years, rested for y years, repeat
- $[(x-y)^3, (a-b)^5]$ = 3 panels with $x-y$ scheme, 5 panels with $a-b$ scheme
- “n” means “never” go back

Therefore,

- $[1-0]$ = always revisit
- $[1-n]$ = never revisit
- $[(a-b), (c-d)]$ = split panel

Scheme 1 can be described as [1-8], Scheme 2 by [2-7], and Scheme 3 by $[(1-0)^1(1-7)^8]$. The advantages and disadvantages of different panel designs are described below:

	[1-0] – always revisit	[1-n] – never revisit	[x-y] & [(a-d)(c-d)] split
Pros	<ul style="list-style-type: none"> • easy to lay out • easy navigation & identification of sample units after one year • most powerful to detect linear trend 	<ul style="list-style-type: none"> • low to zero response burden (wear and tear) • can automatically account for emigration & immigration (each sample is independent of all earlier ones) • efficient estimator of status 	<ul style="list-style-type: none"> • medium response burden • compromise between estimation of status and trends • can be connected
Cons	<ul style="list-style-type: none"> • high response burden placed on sample units (wear and tear) • emigration and immigration of sample units difficult to handle • poor estimator of status 	<ul style="list-style-type: none"> • low efficiency for trend estimation • not connected 	<ul style="list-style-type: none"> • planning more complicated • optimum allocation of effort among panels is complex and depends on the contrast

Designing Panel Surveys for National Parks in the Northwest (presented by N. Scott Urquhart, Colorado State University; this presentation is based substantially on Urquhart et al. 1998; see it for further detail and clarification of his analysis.)

Introduction

Inference from survey data can be accomplished from three perspectives:

- **Design-based** inference is based on estimates derived from the data and applies to an explicit sample frame. Randomness is drawn from the design itself. This provides the clearest link to park-wide inferences, has the least number of assumptions, and would be the type of inference one would want for testifying to Congress. However, it takes no advantage of auxiliary information (e.g., weather patterns for vegetation).
- **Model-assisted** inference uses a model to complement the sampling structure, thereby taking advantage of auxiliary information. For example, if bird populations along trails are considered to be representative of park-wide populations, then data taken near trails can be used to make inferences to the park. However, you must be able to defend the assumption of the model indicating that populations near trails are representative of the park.
- **Model-based** inference ignores the sampling plan altogether and depends on populating a model with the observed data (e.g., spatial statistics). The validity of the conclusions depends on the validity of the model.

Regional Trend versus Site Trend

The predominant theme of ecology is to understand ecological processes (e.g., energy flows, food webs, nutrient cycling) to determine how a particular ecosystem functions. Studies with the ability to answer these questions must be temporally intensive (e.g., many samples from each plot per season), and thus spatially restrictive. Consequently these data cannot provide

precise inference to an entire region, however that is defined. Therefore, the intensive approach is not appropriate for NPS monitoring, which must focus on an entire ecological resource across an area or region, including all of the available variability. ‘Region’ can be defined as individual parks or the entire network.

What is Trend?

Trend is any response that changes across time in a generally increasing or decreasing manner (i.e., change has no direction, but trend does). Even if trend is not linear, it will have a linear component that will be detected as linear trend. Some variables may have a pattern that is not a trend and vice versa. For example, an ascending line has trend but no pattern while a sine curve has pattern but no trend. (These definitions are not used by all statisticians.)

It is nearly impossible to detect trend in less than five years. This is because

$$\text{Variance of slope} = \sigma_{\text{slope}}^2 = \frac{\sigma_{\text{observation}}^2}{\sum (t_i - t_{\text{avg}})^2} \quad \text{where the denominator indicates data points through time.} \quad (1)$$

After two years you have only measured change, after three years, the denominator equals 2, after 5 it equals 10, after 10 it equals 82.5. Consequently, the variance of the trend estimate decreases with time making trend easier to detect.

Patterns of Variability for Indicators

Biological indicators tend to have more variability than physical indicators (e.g., acid neutralizing capacity and conductance are less variable than number of zooplankton or rotifers). This variance has three components: population (spatial), year (temporal), and residual variances.

- **Population (Spatial) variance** is the same as site variance and is the variation among values of an indicator (response) across all sites in a park or group of related parks, that is, across a population or subpopulation of sites. The site component of variance is one of the major descriptors of the regional population.
- **Year (Temporal) variance** is the concordant variation among values of an indicator (response) across years for all sites in a regional population or subpopulation; it is not the variation in an indicator across years at a single site. If trend is present, year variance is effectively the deviation away from the trend line (or curve) after trend is accounted for. The year component of variance is often too small to estimate and it is the enemy for detecting trend over time. If it is even moderately large, the ‘sample size’ reverts to the number of years for which there are data, and the number of visits and/or number of sites have no practical effect on trend detection.
- **Residual variance** results from a year-by-site interaction (i.e., the site-specific part of what ecologists would call year-to-year variation) and index variation. Index variation is due to measurement error, crew-to-crew variation, local spatial variation, and short-term temporal variation (when in the index period the sample is taken). Residual variance characterizes the inherent variation in the response or indicator, but some of its subcomponents may contain useful management information. Crew effects could indicate the need for better

training, visit effects could indicate the need to redefine the sampling period (index period), and measurement error could indicate the need to improve laboratory techniques.

The partitioning of these components of variance for data collected at northeastern lakes indicates that the distribution of variance varies with indicator (Figure 8). Some responses are dominated by residual variation and some have year effect. If the year variance is high, it is difficult to detect trend. The site variance was the largest component for most attributes.

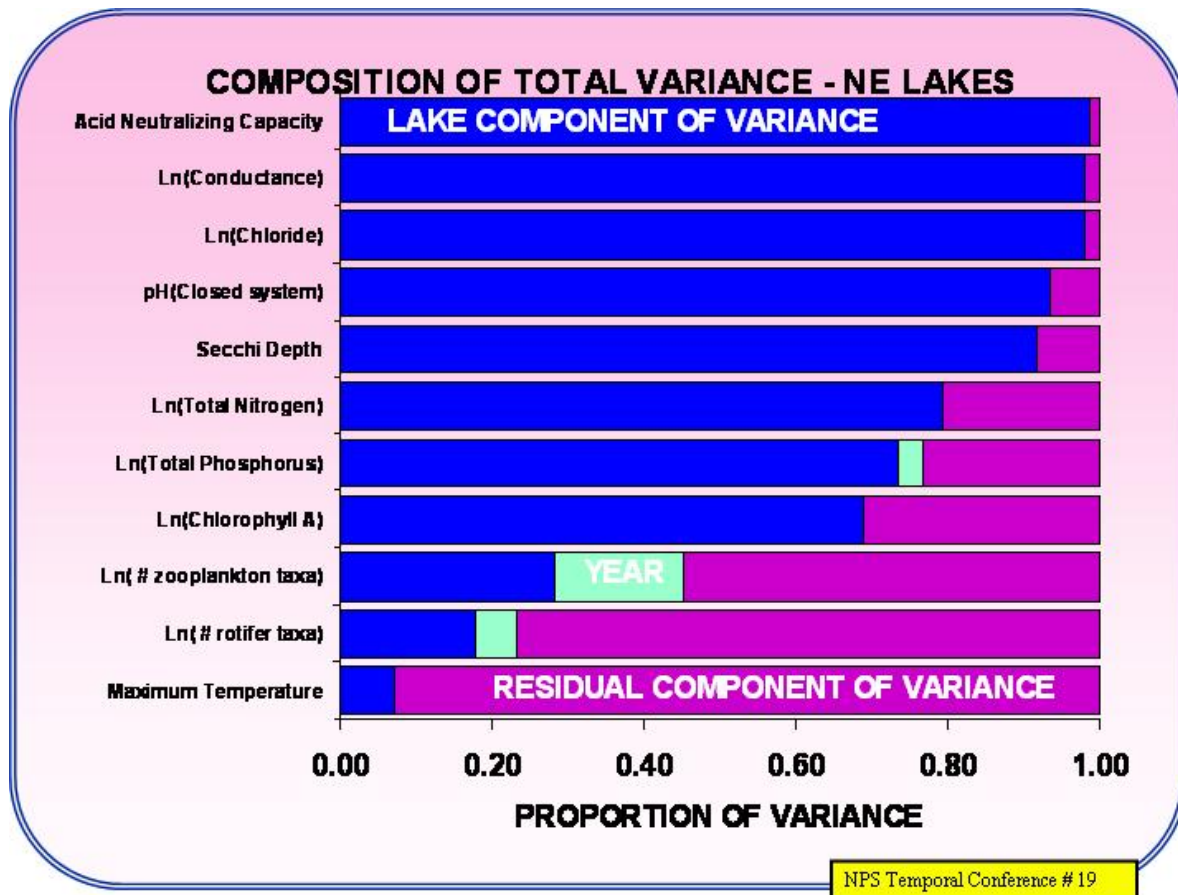


Figure 8. Partitioning of total variance in Northeast lakes (From Scott Urquhart's presentation; lightest shade is the year component of variance).

The Oregon Department of Fisheries and Wildlife stream habitat survey is a good source for the partitioning of variance components for Pacific Northwest streams. Note: ocean effects drive the presence of anadromous fish in streams so covariates describing those effects must be identified before you can understand the variability in the data.

Design Tradeoffs – Trend versus Status

So how do we detect trend in spite of all of this variation? We rely on the statistical properties of the variance of a mean, and blocking.

$$\text{Variance of a mean} = \frac{\sigma^2_{\text{slope}}}{m} \quad \text{where } m \text{ members of a population have been randomly selected and their responses averaged.} \quad (2)$$

This variance can be reduced by increasing the number of sites. Remember that the ‘mean’ is a regional average of slope (the mean of trends), so the numerator refers to the variance of an estimated slope. Then the variance of the mean slope becomes (combine equations (1) and (2)):

$$\text{Variance of regional mean slope} = \frac{1}{m} \frac{\sigma^2_{\text{observation}}}{\sum (t_i - t_{\text{avg}})^2} \quad (3)$$

Therefore, the variance can be reduced by increasing the number of sites (m) or the number of years (t). Note that the regional averaging of slopes (m effect) has the same effect as continuing to monitor at one site for a much longer time (t effect), but there is more benefit from increasing the number of sites (m).

The total variance of trend estimates in a sample is large. If we take one regional sample of sites at one time and another independent sample at another time, the site component of variance will be included in the variance of each sample, but we will have a better spatial distribution to give an idea of status. However, if we consider each site to be a ‘block’ and periodically revisit a site, the site component of variance disappears from the variance of a slope. In general, trend detection improves with more revisits to sites and status estimation improves as the number of distinct sites increases. The NPS monitoring program will likely be based on a constant effort per year. Therefore the effort will have to be distributed to optimally balance the trade-off between the total number of sites sampled (to estimate status) versus the number of revisits to sites (to estimate trend).

Recommendations for NCCN Parks – Temporal Design

What kind of temporal design should NCCN Parks use? Realistically, the NPS monitoring program will be based on equal effort per year. Therefore, we will investigate two families of recommended designs based on 30 site visits per year. General conclusions from these examples apply to all sample sizes, but higher sample numbers will have better specific performance. We will use the notation presented by Trend McDonald above.

The first family of temporal designs (DF1) consists of panels of [1-0] and [1-n] (Figure 9) with several distributions of the 30 site visits to the panels. We will consider the following distributions of site visits between the two panels:

[1-0]	30	20	10	0
[1-n]	0	10	20	30

The first design is equivalent to ‘always revisit’; the last design is equivalent to ‘never revisit’.

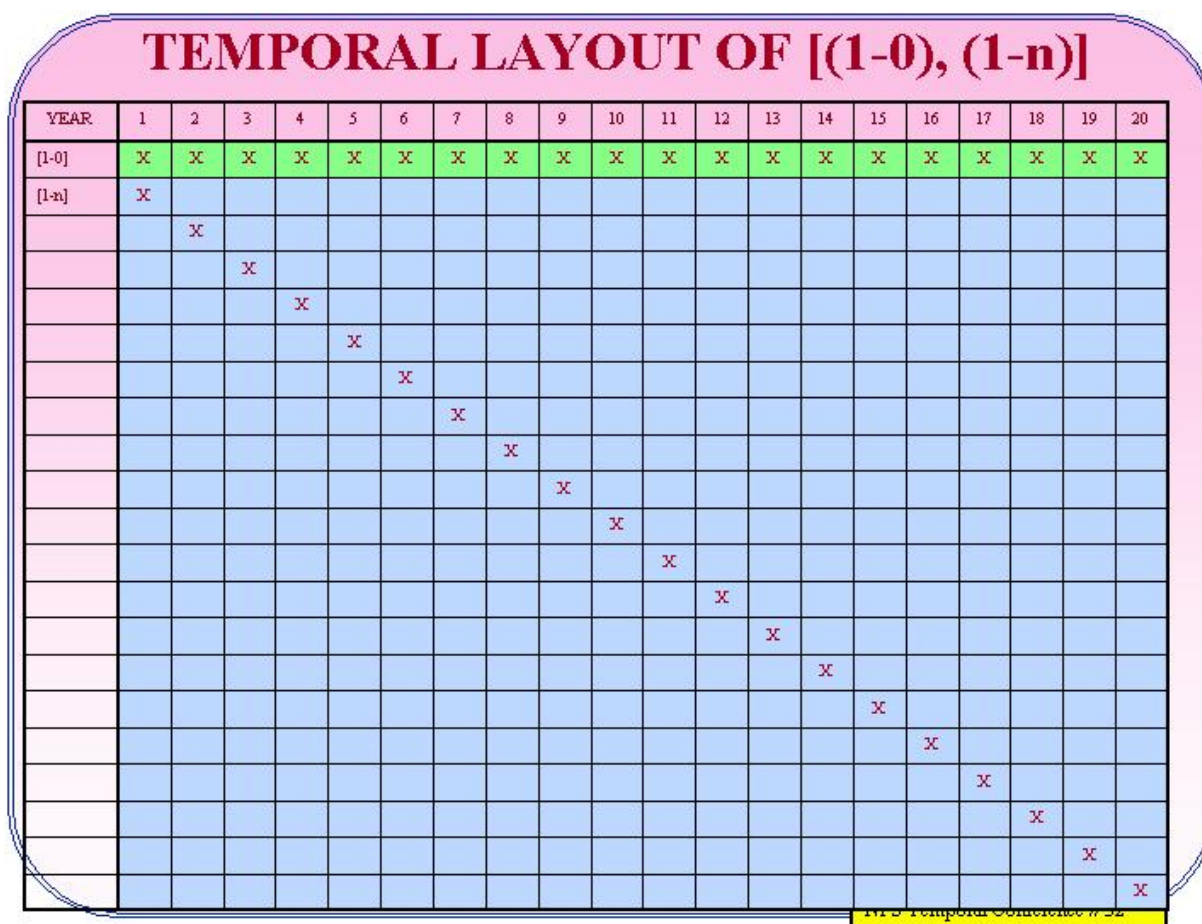


Figure 9. Design Family 1 (From Scott Urquhart’s presentation).

The second family of designs (DF2) is called ‘split panel’ designs and has the feature of no panels that are visited every year. We will vary x and y in the $[(1-4)^x(2-3)^y]$ design as an example (Figure 10). These designs have the very desirable feature of being ‘connected’ in the experimental design sense because some plots are measured in consecutive years. Connectivity provides the possibility to estimate year effects, if present, and is important if the objectives include estimating annual means and differences/contrasts among them. It is not necessary if the objective is to estimate trend as an average rate of change (e.g., slope of a trend line). We will consider the following distribution of site visits between the two panels:

[1-4]	30	20	10	0
[2-3]	0	5	10	15

Note that when sites are visited in two consecutive years, it is only possible to visit half as many sites compared with only visiting them once.

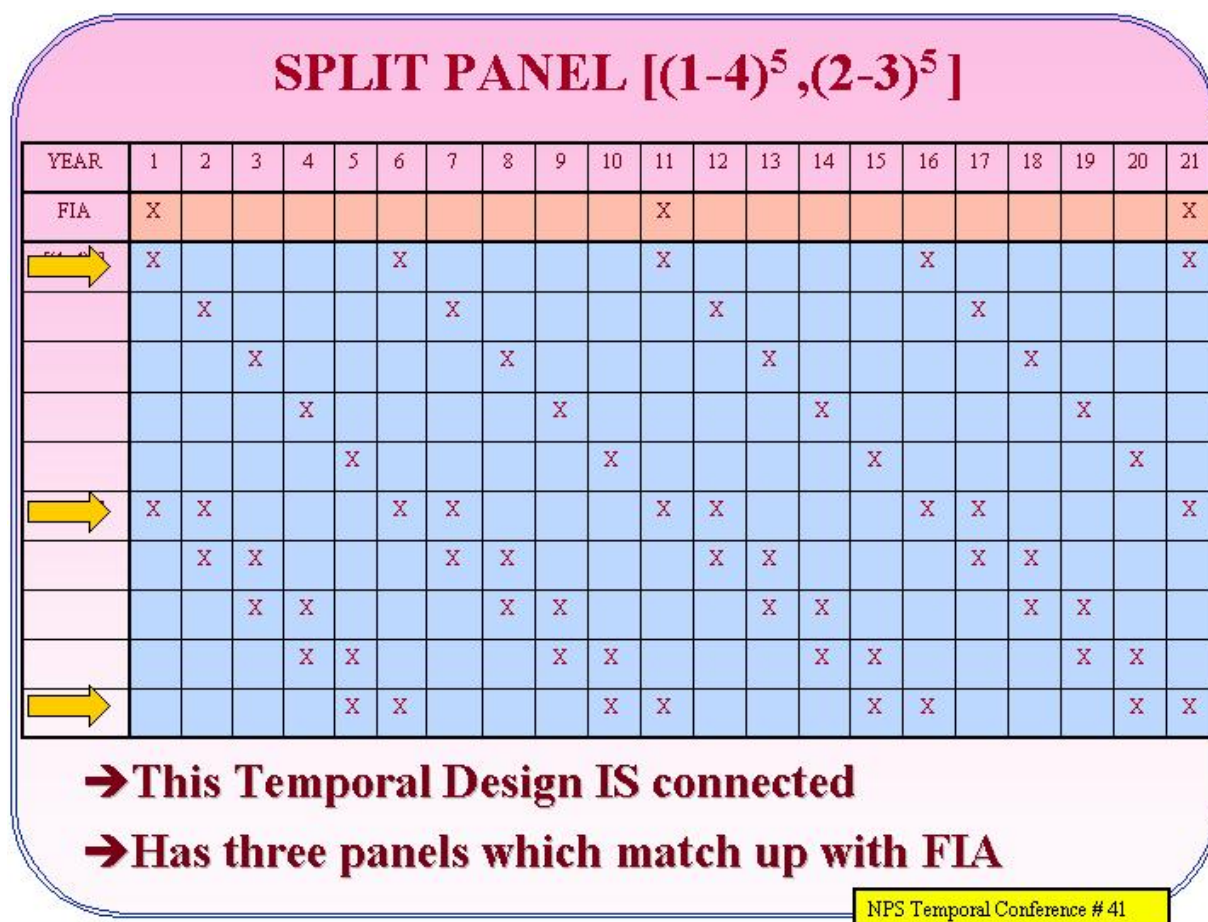


Figure 10. Design Family 2 (from Scott Urquhart’s presentation; ignore FIA line).

Power to detect linear trend from the slope of the trend line was evaluated within and between the design families while keeping site number and variation among sites constant. This analysis for DF1 shows that the 'always revisit' design always has higher power than the other three. However, the difference among designs and the power of all designs decreases with increasing year effects if they are not modeled (Figure 11). The power analysis of DF2 shows very little difference among individual designs (Figure 11) and the power is equivalent to the best power of DF1 for the same amount of year effect. Also, the power to detect trend increases faster with time for DF2 versus DF1.

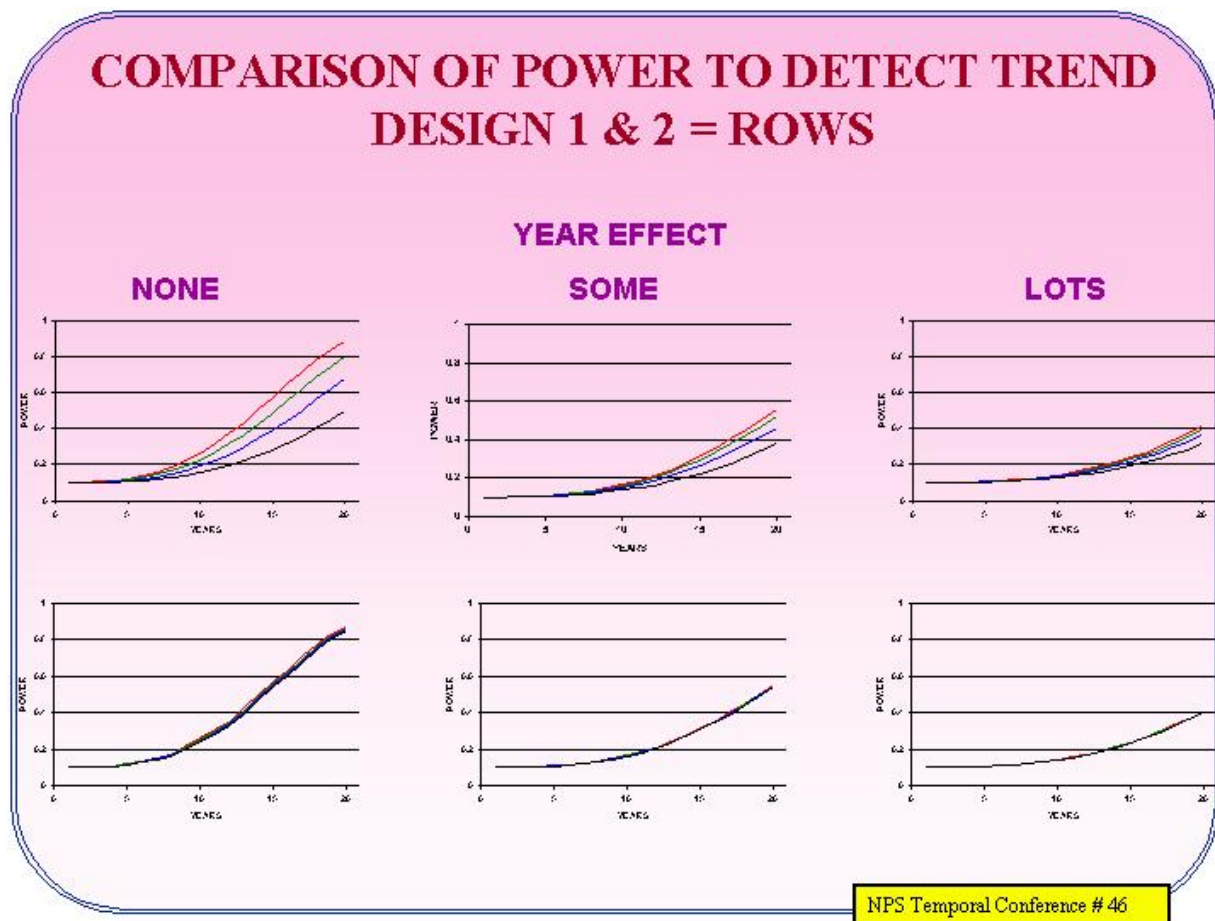


Figure 11. Comparison between DF1 (Various distributions of 30 plots between (1-0) and (1-n) panels) and DF2 (various distributions of 30 plots between (1-4) and (2-3) panels; split panel design) for ability to detect trend given different amounts of year effect (from Scott Urquhart's presentation). Row 1 illustrates increasing year effect for DF1; row 2 is DF2. Lines in Row 1 from top to bottom are 30:0, 20:10, 10:20, 0:30 distributions of plots between [1-0] and [1-n] panels.

An analysis of the standard error of the status estimate for each design shows that ‘always revisit’ always has the highest standard error among the DF1 designs after the first few years, the others are not very different from each other, and the standard error increases for all with increasing year effect (Figure 12). DF2 has very little difference among individual designs, especially in the early years of monitoring (Figure 12). DF2 all start with similar standard errors to DF1, but they all improve over time similar to all but the ‘always revisit’ version of DF1. The kink in the graphs for DF2 where the standard error decreases abruptly occurs when sites are revisited for the first time.

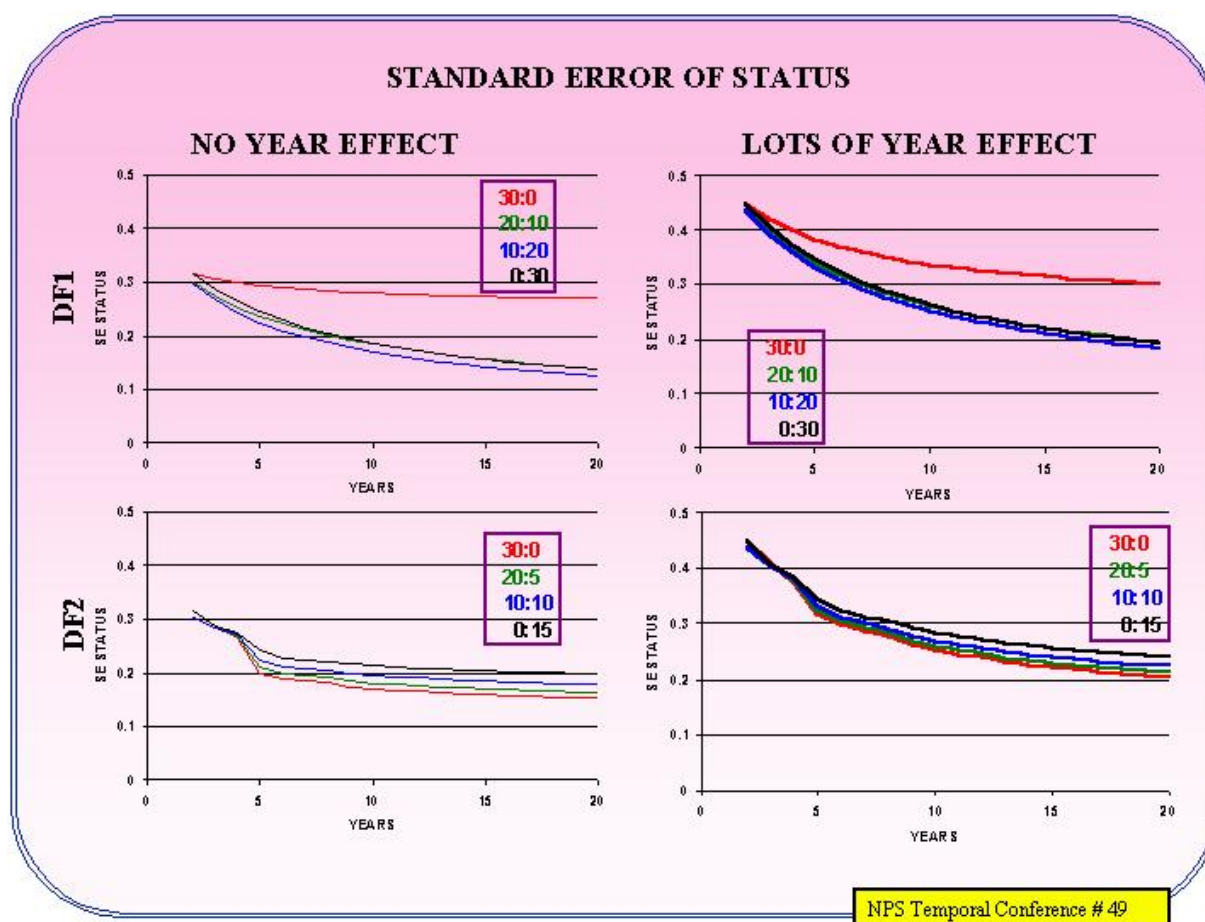


Figure 12. Comparison between DF1 and DF2 to describe status for different amounts of year effect (constructed from Scott Urquhart’s presentation; see Figure 11 for definitions of DF1 and DF2).

These examples demonstrate that trend detection improves with the number of revisits, and status description improves with the number of sites. Between the two families of designs, the power to detect trend in DF2 is equivalent to the best power in DF1 and it shows better improvement with time. DF2 also has equivalently low standard errors for status as the best of DF1. Notice that the proportion of panels allocated to the revisit design only slightly influences power to detect trend. Consequently a split panel design similar to DF2 is recommended for the

NPS. Another example comparison from a range of panel designs is shown in Appendix B including varying the effect of sample size and the size of the trend.

For presentations of data sets, the information given by the presenter will be recounted, followed by a summary of the discussion. The group did not support all ideas given in the original presentation; therefore a summary of the recommended approach reached by the end of the discussion is also given. The goal is to present the original concepts of the presenter and how they were modified by the ensuing discussion.

Developing a Sample Design for Avian Monitoring in the NCCN (presented by Kurt Jenkins, USGS-FRESC Olympic Field Station)

Background

The USGS Forest and Rangeland Ecosystem Science Center and the Institute for Bird Populations (IBP) are working with NPS to design and develop protocols for monitoring changes in the abundance of breeding landbirds in the NCCN. The IBP has recently completed two-year inventories of avian abundances and distribution patterns in North Cascades and Olympic National Parks and will complete a second year of inventory in Mount Rainier National Park this summer. In 2002, the USGS and IBP entered into a cooperative agreement to develop peer-reviewed avian protocols for the NCCN.

Monitoring of birds is important because NCCN parks are the largest un-harvested blocks of habitat in the Pacific Northwest, and they are the best place to study bird dynamics to tease out effects of land-use outside of National Parks versus changes in tropical habitat. Monitoring in national parks is desired by USFWS because many species common in parks are of national importance.

The objective for this session of the workshop is to recommend a sampling design for avian monitoring in the NCCN. While the emphasis is on recommending a temporal sampling frame, we will first review the spatial sampling designs as background. We will focus our discussion on Olympic National Park as a prototype, representing the sampling issues of other large parks.

An important step in developing this avian monitoring program is a power analysis conducted on three years of pilot data. The analysis determined the relationship between the number of transects and power to detect a positive or negative 4% annual exponential trend in average daily counts of birds for 22 species. The analysis assumed that data would be collected annually and analyzed using a linear regression profile summary. This analysis reflects how the Bird Workgroup expects to analyze the data and what they plan to measure. It suggests that large numbers of transects distributed among diverse vegetation types are needed to effectively detect trends in as many species as possible.

A meeting of avian monitoring specialists in NOCA in September 2000 provided initial direction for developing avian monitoring in the NCCN (Siegel and Kuntz 2000). The group (mostly) agreed that the goals of avian inventory and monitoring should be to elucidate

- spatial patterns of abundance from broad-scale monitoring
- **temporal patterns of abundance, also from broad-scale monitoring**
- demography of target species from intensive monitoring, although this goal is probably not financially feasible.

We will focus on the second objective, recognizing that the spatial component was largely completed in the inventory phase. We agreed that the monitoring program should provide information on temporal trends relevant to park management issues at individual parks but should also elucidate regional trends on these, the largest protected lands in the Pacific Northwest. Therefore, our objective is to sample each of the three large natural-area parks sufficiently to derive park-specific inference on population trends of common avian taxa. We plan to use identical sampling strategies and methods to permit the pooling of data for regional trend analyses. Monitoring should be sufficient to detect at least a 3% annual decline of salient species over 20 years (45% decline over 20 years) or 4% annual decline of salient species over 12 years (a 39% decline over 12 years). These are somewhat arbitrary numerical objectives being used as a starting point for developing preliminary sampling guidelines.

We will sample avian detections from clusters of generally 5-12 points distributed 200-m apart along independent transects (i.e., primary sampling units). Changes in avian abundance will be inferred from two primary metrics: 1) We will estimate changes in density of the more common species, for which we are able to derive detectability functions, using distance-based sampling theory and estimation programs (Program DISTANCE 4.1). 2) To complement changes in density we will also infer changes in abundance from changes in the proportions of sample points where species are detected (Program PRESENCE). Program PRESENCE requires multiple site visits to at least a sample of transects for computation of detection bias. The response design consists of detecting birdcalls in variable circular plots (VCP) centered on points evenly spaced along transects. The observer listens from the point and the detection probability of a call is a function of the bird's distance from the point. Detectability can vary greatly among species because some are more/less easily heard. Distance sampling theory (Buckland et al. 2001) is used to estimate density from the raw counts and the effective area sampled (derived from the detectability function). Observers must be at the starting point at 5:00 am and continue sampling until 4 hours after sunrise.

Constraints

Olympic, like the other large parks in the NCCN, presents several real challenges for sampling. Much of the park is remote, steep, wet, slippery, or crumbly. Further, fluctuating river levels prevent reliable safe access to many points. Slipping, falling, and carrying excessive weights are common sources of injury for Natural Resources employees of these parks. For monitoring to be sustainable, sampling must be safe and reasonable. We routinely rule out sampling on slopes $>35^\circ$ and we often subjectively rule out other points as not safely accessible.

The greatest level of support we can reasonably expect for avian monitoring in Olympic National Park is a field crew of 4 technicians working for about 8 weeks each summer (this excludes time needed for training, planning, and data management activities each year), and half of that level of support is a very real possibility. It is difficult to gauge the rate of sampling that can be accomplished with that effort because it depends so heavily on sampling design.

But it is useful to consider that we might sample about 64 transects in a season if each sampling transect were accessible within a half a day of travel from the previous transect (i.e., all sites are road accessible or closely clustered in the backcountry). By contrast, we could probably sample 16-32 transects in a season if plots were randomly or systematically distributed throughout the park (i.e., 1-2 plots sampled per week due to travel costs).

Generic Sampling Design

We have developed general sampling recommendations for long-term monitoring projects in Olympic National Park (Jenkins et al. 2002). We recommended stratifying the park according to three categories of accessibility and human use to promote flexibility in delineating the target population and varying sampling probabilities in relation to access costs (Figure 13):

- High Accessibility/Human Use: Areas <1.5 km from a maintained park road
- Moderate Accessibility/Human Use: Areas <1.5 km from a maintained hiking trail.
- Low Accessibility/Human Use: Areas >1.5 km from a maintained road or trail

In addition to the flexibility of such a design in the allocation of sampling effort, these strata also partition gradients of human use, one of the notable disturbances suspected of causing change in park resources. On occasion, a tremendous effort is required to hike >1.5km from a maintained trail and suitable camping areas are often lacking away from trails. Because bird crews must begin work at dawn, we propose defining the target population for avian monitoring in Olympic National Park as areas within high and moderate accessibility/human use zones (subject to the safety filters). We may want to consider also adding a sampling stratum to include monitoring from the actual trail. While such a sampling strategy would limit inference to approximately 50% of the park, all vegetation types are included within this sampled area.

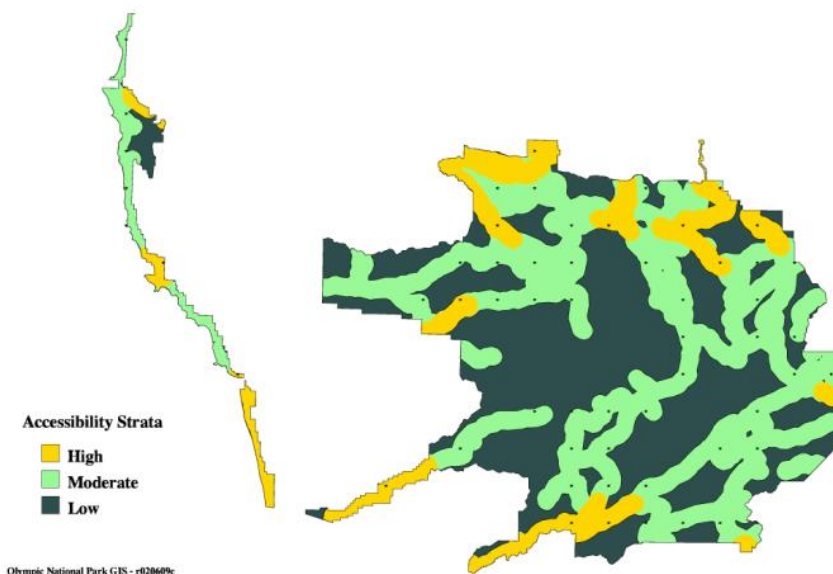


Figure 13. Example of a systematic sample of transect starting points within high and moderate accessibility strata in Olympic National Park (Jenkins et al. 2002).

We propose to sample 208 point count transects in Olympic based on the budget and the experience of IBP. Each transect has points every 200 m where the observer stops for 5 minutes and records detected bird calls by species. Eighty-two transects are off-trail from systematic sample points and 41 are on-trail. We will focus on the 41 backcountry transects as a systematic sample (meaning we will sample a lot of western hemlock and Douglas-fir communities because they are common).

Discussion Points

The discussion of bird monitoring was wide-ranging and occurred over 3 days. The following summary is presented by discussion topic, rather than chronologically, to enhance clarity. The discussion was mainly between the invited statisticians and the Bird Workgroup with Kurt Jenkins moderating.

Need for better objectives

In the background presentation for birds, Kurt described the objectives of this monitoring project as detecting trends in spatial and temporal patterns of abundance of common bird taxa with park-specific inference to high and moderate access areas. The goal is to detect a 3% decline over 20 years or a 4% decline over 12 years. Abundance would be estimated using variable size plots (VSP) and the program DISTANCE as well as by proportion of sites occupied (PSO) and the program PRESENCE. After some discussion it was decided that these two metrics are redundant and noted that the second one requires repeat visits. Consequently, the Bird Workgroup agreed to focus on VSP.

The statisticians were not satisfied with the detail of the stated objective. After additional discussion, the Bird Workgroup revised the avian monitoring objective as follows: Determine long-term trends in species composition and density of landbirds in accessible areas of NCCN parks during the breeding season. The target population was specified as common breeding terrestrial song birds, especially the approximately 20 species for which power analysis indicates trends can be detected. The goal is to detect small annual changes over decadal time periods, although year effects are also important. The metric for species composition is yet to be determined, and the density metric is birds/ha/species without regard for habitat. However, there is interest in distributing sampling spatially to ensure a reasonable sample of habitats to maximize the number of species detected.

Riparian areas were identified as potentially troublesome habitats because they are extremely important with respect to bird diversity, but they are poorly sampled by a 2-D approach. They may require a separate linear sample in riparian corridors.

Range of Inference: Park versus Network

An alternative to achieving park-specific inference would be to obtain network-wide inference. This approach would take advantage of the greater ease of reaching some ecosystems in some parks than others (e.g., subalpine and alpine areas are much more accessible in MORA than the other large parks). All ecosystems could be sampled somewhere in the network, but not

necessarily comprehensively at any park. This approach makes sense because bird populations are a regional issue and the three large parks include many of the same species.

This argument was countered by the recognition that the audience for monitoring is park managers and the public, who are interested in park-specific summaries. Also, this approach implies that a region in one park (e.g., subalpine) represents that region in all other parks, ignoring the inherent variability among parks. Nevertheless, observations in other parks provide a needed context for each park and contributes to ‘weight of evidence’ arguments to support regional conclusions. It was decided to continue with park-wide inference. In fact, both objectives are achievable with the same data, with the network estimates being more precise. For the network objective, you would want to optimize the effort among parks, putting more effort in parks with the easiest access. For the park objective, you may want to put equal effort or an equal number of sample points in each park.

Range of Inference: Accessible Areas versus Entire Park

Cons for Sampling Only Accessible Areas. The inability of the proposed monitoring program to say anything about inaccessible areas was troubling to the group. Participants felt that the park would be paying a big price for the future by not monitoring the entire park. Short term consequences are that you can say nothing about ONP, for example, only about its accessible areas, thereby compromising the goal of NPS monitoring to describe entire parks. In the long term, such a monitoring program would not be able to answer questions not yet imagined.

Although Rodney’s experience indicates that bird populations are not different one-half mile from the trail compared with three miles, we don’t actually know how bird populations change from one accessibility stratum to another or how far away from a trail ‘normal’ behavior occurs. In fact, detectability may even be lower off-trail compared with on-trail due to the disturbance created by observers crashing through the brush. On the other hand, reliability doesn’t necessarily increase by staying in accessible areas and it may decrease if birds avoid trail corridors. It would be very interesting to test Rodney’s hypothesis because if the difference is small, it would make the on-trail estimates more reasonable. If the difference is large, you can use the off-trail point to make double sampling estimates, which corrects the on-trail estimates using the ratio of off-trail to on-trail.

There are also technical problems with defining ‘accessible’ and park staff may not be able to accurately map inaccessible areas. Experience shows that even when an area is determined to be accessible by a GIS analysis, in fact the area is not accessible when investigated on the ground. Even if it could be mapped now, accessibility may change in 10-20 years and the fixed strata boundaries would no longer reflect reality. Also, accessibility is worst for the first visit. After that, better mapping of access trails and local notes make points easier to find. Moreover, only part of the annual sample will be new after the first year if using a connected design.

Pros for Sampling Only Accessible Areas. Kurt Jenkins and Rodney Siegel presented strong arguments for the need to limit sampling to accessible areas of the park, despite the resulting reduction of range of inference. Their arguments focused on feasibility and safety of field crews. A primary constraint is the need for crews to be at the transects in the dark (at 5

am) in places that may have steep terrain and are unsuitable for camping due to lack of water. If tremendous time, money and injuries are expended to collect relatively few data in inaccessible areas, park managers are less likely to sustain support for the program. It is also possible to lose field crew support if their willingness to get to difficult places is over-taxed. The need for park-wide inference may be overstated because there are successful monitoring programs in other parks that are based entirely on trails (e.g., Great Smokey Mountain National Park) or roads (i.e., Breeding Bird Survey). While NPS guidance charges managers with managing the entire park, there should be exceptions for large parks with inaccessible areas. However, this constraint should not hold for small parks.

In the introduction, Kurt pointed out that strata reflect human use to some extent. This was mentioned as a potential strength of stratification by access. Upon reflection, the group realized that the strata do not accurately represent zones of human influence, which are much more restricted to trails and campsites. The discussion ended with an explanation that an objective of monitoring was not to monitor the effects of human activities, and that accessibility strata reflect only the logistic cost of conducting sampling.

Conclusions and Proposed Solutions. The group accepted that NPS monitoring must make serious compromises for reasons of safety and cost. The monitoring program must balance the desire for inference against the other needs competing for the same funds. All agreed that limited sample effort at this junction is especially warranted because birds are not likely to be a hot policy issue or a long-term regulatory issue. However, caution was advised not to tie the hands of future researchers by design-related decisions. One alternative is to recommend an overall design using the entire park, but limit initial implementation to accessible strata.

Further discussion focused on block designs based on watersheds to reduce travel time and increase sample number. The merits of interpenetrating panel designs (distribute sample throughout park) and block designs (sample a subset of watersheds each year) relate primarily to sample effort. Block designs allow for better resolution of between year effects, but they also inseparably confound watershed and year effects. For example, perhaps 10 sites could be sampled in a single watershed in 10 days, while 10 interpenetrating samples may take 30-40 days to sample. Therefore, a panel design using watershed to assign membership will allow a greater number of total sites, but it may take 4-5 years to complete a rotation of all watersheds. The watershed approach also has the accessibility issue, but has the advantage of grouping points that might be done in a two-week stint. One could also define clusters as groups of transects accessible in 10-days. This is based on the usual field schedule for the crew (7 days on, 3 days off). So, two eight-week crews could sample 16 clusters per year. Some revisits need to be built into the design and could be accomplished by changing cluster composition between years with some of the sites re-measured from the previous year to develop connectivity. Finally, a hybrid approach was proposed to put the trails in one panel and sample them annually (either exhaustively or some annual subsample), then allocate additional samples among watersheds at a less frequent interval. In the final analysis, the group chose at least a partially interpenetrating design with panels consisting of nearby plots.

Spatial Design: How to Distribute Samples to Meet Multiple Goals?

Problem. We have discussed the need to distribute sample points on the landscape to reflect their accessibility. Meanwhile, a biologically balanced sample is also needed. We would like to use strata (i.e., geology, topography, watersheds) to determine biologically meaningful space (as opposed to accessibility strata) and to distribute our samples among them. Divisions that have small area can be adequately sampled by raising the sample intensity in that area or by raising the inclusion probability. For example, approximately 15-25% of OLYM is above treeline. To be sure that this portion of the population is included we may give the higher elevations a higher selection probability or allocate sample numbers to vegetation categories. The need is to define strata and decide how many samples out of the total should be in each stratum. One method would be to base strata on partial lists of bird species with the intent to have all of the relevant species well sampled. Kurt clarified that the goal of the bird monitoring program is not to sample specific bird communities. Instead the goal is to adequately sample all habitats to maximize the number of species encountered. Therefore stratification should be based on physiographic boundaries approximating vegetation zones rather than habitat delineations based on bird distributions. The group of experts was generally uncomfortable with defining strata using sharp boundaries as proposed in Kurt's introduction. They felt that the boundaries could not be accurately mapped and that the true boundaries would vary in the future. (Note: selection probability is the probability that a point will be chosen for the sample at each draw; inclusion probability is the probability that a plot will be included in the sample after all points have been drawn.)

This argument was countered by the observation that an accurate map is not needed to effectively allocate samples spatially. Any strata will achieve that goal and will provide unbiased estimates of parameters. While the boundaries of strata are fixed forever, there is flexibility in the sample size assigned to strata through time, and strata can be split into smaller units in the future. Therefore a stratified spatial sample is not completely rigid.

Solution. The proposed solution to achieving multiple goals (e.g., accounting for access and physiography) while distributing sample points was to use variable probability sampling. Variable probability sampling offers an alternative to stratification that accomplishes the same results (i.e., allows you to distribute samples according to multiple priorities). The method involves using GIS to establish a fine grid of points over a park. These points would be overlain with a probability surface based on safety and cost, and perhaps a second surface that gives a probability inverse to size so that scarce habitats (elevation zones) receive a higher probability of being sampled. The result is a grid of tight points, each of which has an associated probability of being sampled equal to the product of the probabilities identified in the layers that overlay it. Sample intensity can be unequal among biological strata. For example, if areas are quite small and/or significant, their probability can be increased to ensure they are captured within the sample. For bird monitoring in NCCN the important factors are riparian areas, for which certain species show affiliation, and elevation as a surrogate for vegetation pattern and accessibility. A grid of points, each of which has a cost/safety and elevation factor, will produce a response surface of sampling probability (probability of unsafe areas could be set to zero). This surface provides the foundation for spatially distributing the sample. An excellent example is a study of the California bight (Stevens 1997; see also Woodward and Jenkins 2001).

Two cautions for creating variable probability samples were discussed. First, sample size is a random variable, meaning that the sample size in any stratum will vary around the intended number. This could be a problem if sample sizes are reasonably small because by chance you may have insufficient samples in some strata and over-sampling in others. There was also discussion that the range of probabilities be truncated so that there isn't more than a 1:3 to 1:5 ratio between the highest and lowest probabilities. Otherwise you may happen to select a sample that has a very low selection probability and is very different from other sites. This site can have a huge impact on the regional trend results because its low probability will cause it to be weighted heavily in the final analysis. This is also a problem if you have a large stratum with a small sample size. You are more likely to sample an unusual site with a small sample size and it will inaccurately represent a large proportion of the entire park and will significantly increase variance.

Temporal Design: Status versus Trend

The discussion of an appropriate temporal sampling design began by noting that the appropriate revisit interval depends on the organism to be monitored. For example, temporal patterns of variation for sedentary, long-lived vegetation are not the same as those for mobile, annually reproducing birds. Also, vegetation permanent plots are much more subject to fatigue due to trampling than birds plots. Finally, the advantages of revisiting sample sites is not degraded by the inability to precisely relocate the sample site in cases where site affinity is low (e.g., birds show weaker affinity than plants). In conclusion, vegetation plots need not and should not be monitored annually while it is possible, and necessary for detecting intra-decadal cycles, to monitor birds annually.

As illustrated in the introductory presentations, trend is best detected with annual revisits to plots, while status requires the greatest possible spatial coverage. If as stated, the focus is on trend, then the optimal sample design utilizes annual visitation of fixed plots given trampling is not a problem. However, this would greatly reduce the spatial extent of the sample and it is important to encounter as many species as possible. A split-panel design (i.e., one that includes revisits at some interval) is most preferable because it allows greater spatial coverage with only marginal reduction in the power to detect trends. Therefore, the best design would have some panels that are revisited annually and others that may be revisited at 4- or 5-year intervals. This is a compromise between the two objectives of describing status and trends. If one objective is more important, you could achieve greater precision by concentrating on one objective and not diluting your effort by trying to do too much.

Possible Panel Designs

The Bird Workgroup initially suggested a $[(1-0), (1-3)^4]$ design where each panel consists of 20 transects broken into 4 field tours of 5 transects each for a total of 100 transects. One panel of 20 transects is visited annually and 4 panels of 20 transects each are sampled every 4th year. It was noted that a $[(1-0), (1-5)^6]$ design would take longer to detect trends but would provide greater spatial coverage. If each panel consisted of 20 transects, a total of 140 transects would be sampled over a 6-year rotation. Another option is panels that are sampled in 2 consecutive years followed by some interval. There was a consensus that a small number of (1-0) samples should be completed (this assumes that plots do not wear out), however, the proportion of effort among annually revisited transects and other, less frequently sample sites was less clear.

Some argued that spatial coverage is important (i.e., fewer annually monitored sites). A good compromise among trend, status and connectivity may include a (1-0) panel of 15-20 transects (3-4 field tours of 5 transects) where field tours represent the major subdivisions of the park, plus (1-4)⁵ panels of 20 transects.

	Year											
Panel	1	2	3	4	5	6	7	8	9	10	11...	
1	20	20	20	20	20	20	20	20	20	20	20	
1	20					20					20	
2		20					20					
3			20					20				
4				20					20			
5					20					20		

Membership Design

It is much more efficient for the Bird Workgroup to put sites into clusters of ‘field tours’ within spatially connected areas so that one crew can complete four field tours in one season. Kurt asked whether they should select transect starting points or field tours first. The majority of specialists recommended that the points be selected first, and then assigned to field tours. If you assign the points to the field tours prior to selecting which points will be sampled, then you are doing cluster sampling. However, you CAN assign the sites (field tours) to panels after they are selected. You would want to spread the 8 field tours measured in any given year both spatially and according to elevation.

Paul Geissler countered that more efficient field operations can be achieved by selecting tours first instead of assigning pre-selected points to tours. If you assign pre-selected points to tours, you will probably find that some points do not fit nicely into your tours. If tours are selected first, points can be spaced at convenient and predictable distances. This approach amounts to a cluster sample. A cluster sample can be either more or less precise than a simple random sample, depending on the intraclass correlation. For example, a systematic sample is a type of cluster sample that is more precise than a simple random sample because the interclass correlation is negative in the presence of spatial autocorrelation. He suggests that the tours be laid out as a systematic sample, with the points separated by the distance a crew can reasonably cover in a day. Separating the points as much as possible will improve precision by reducing spatial autocorrelation.

Response Design Note

Sample effort is determined by time available in the morning, which results in unequal samples per transect. One option is to standardize plots per transect (perhaps 5), however this results in wasted field time if more than 5 could be sampled. An alternative argument is that the number of plots per site can vary, because the number of subplots is collapsed into a single value to represent each transect. However, you need to maintain a similar number of sites per transect over time. In the end, the number of plots per transect will influence the variance in the estimate of bird abundance at that transect and will need to be accounted for using a weighted or transformed analysis when making larger inferences from the site.

Conclusions in Brief

Spatial Design. Sample sites should be chosen using variable probability with layers for safety/cost and elevation. The issue of sampling riparian areas was not resolved.

Temporal Design. Some sites should be visited annually, and some less frequently (4-5 year interval). The proportion of the effort dedicated to annual versus interval sampling was not clearly decided.

Membership Design. Tours should be spread over the entire park every year to achieve an interpenetrating design instead of a block design based on watersheds. The sample points should be chosen, grouped into feasible tours, and finally the tours should be allocated to panels. Note: After attempting this approach using GIS, the Bird Workgroup decided to adopt Paul's method of assigning tours first, then systematically spacing sample points at convenient travel intervals along the tour.

Designing Monitoring of Forested Vegetation for the NCCN (presented by Andrea Woodward, USGS-FRESC Olympic Field Station)

Background

The USGS Forest and Rangeland Ecosystem Science Center has been working with NCCN to develop monitoring for forest vegetation. Work to date has included three years of pilot data collected in Olympic National Park, followed by power analysis for understory and shrub species. The Vegetation Workgroup of NCCN has not settled on objectives as firmly as the Bird Workgroup. Consequently, these ideas about monitoring forested vegetation are Andrea's, were developed before the advent of park networks, and do not reflect the thinking of the entire group. They were presented at this workshop as a basis for discussing vegetation sample design issues, which will be faced by the group in the near future.

Vegetation is a fundamental part of ecosystems, integrating and reflecting abiotic conditions. Vegetation communities are diverse in NCCN with changes over small spatial scales due to the steep gradients of precipitation and elevation. In addition to indicating environmental changes, vegetation provides habitat (shelter, cover, food) for other species and inputs to aquatic systems. Consequently, changes in vegetation composition and structure signal biologically significant changes in abiotic factors, notably climate, atmosphere, and disturbance patterns, which will have important consequences for other taxa and ecosystems. Although NPS expects to monitor structure and composition of all vegetation layers, in this workshop *we will focus on cover of herbs and shrubs*. Understory species are thought to be more responsive to climate change because their pattern occurs at a finer spatial grain than for canopy species (i.e., understory species indicate the wetter and drier ends of the environmental gradient within a class of overstory). We are interested in measuring cover of dominant, relatively common species that define a plant association group (i.e., indicator species), and to detect change on a decadal temporal scale.

Challenges to monitoring vegetation are numerous:

- High annual variability due to weather and observer effects, obscuring changes due to climate and atmosphere.
- Vegetation is slow to show detectable change except that due to catastrophic events.
- Meaningful interpretation of results must be based on categories of vegetation (e.g., plant association groups; PAGs)
- We cannot distribute the sample based on plant associations because they are expected to change and they may change in ways that cause the old strata boundaries to be meaningless.
- Systematic sampling over-represents common plant associations.
- Vegetation plots are extremely susceptible to response burden (wearing out).
- We have no accurate map of vegetation at the PAG scale.

Monitoring Methods. We plan to monitor vegetation in permanently marked plots. We expect to adopt a plot design similar to the USDA Forest Service FIA plots, which includes subplots nested within a larger 1 ha plot. The FIA sample frame is a systematic sample (~ 5 km between sample points) numbering 123 in Olympic National Park, 93 in North Cascades National Park, and ~ 45 in Mount Rainier National Park. The plots are visited in an unconnected panel design of 10% of the plots each year repeated every 10 years (i.e., $(1,9)^{10}$). We would like to take the greatest possible advantage of these plots in our design, while recognizing that their objectives are different, and may be irreconcilable, with ours.

Spatial Sample Frame. Originally, we planned to use the spatial sample frame concept described in the bird project, where large parks are stratified into access zones (high, moderate and low accessibility) and sampled with a systematic grid sample. Following the bird monitoring discussion, it makes more sense to create a variable probability sample based on precipitation, elevation, and cost of access. The FIA program will potentially be a great asset in the low accessibility zone. Financial resources will determine how much of a park we can infer to, and we may have to limit samples to areas thought to be most sensitive.

Changes in vegetation are usually interpreted in terms of vegetation classes (e.g., plant associations) and the impulse of biologists is to stratify their sample based on the biologically relevant categories. However, given that we want to describe changes in vegetation, it does not make sense to stratify on the classes that we expect to change. So we plan to extract plots in 'domains' from the larger sample for interpretation. In a systematic sample we will likely have 'waste' plots that do not exactly fit into any domain. They may show important changes, but without sufficient replication we will not be able to draw defensible conclusions from them. We also expect to lose plots to catastrophic events. Given the limited funds available, we may need to stratify on elevation or some other fixed characteristic of landscape to obtain enough samples in each domain to detect trends.

Power analysis conducted on pilot data indicate that 7-9 plots have 80% power to detect a 4% annual exponential trend with alpha equal to 0.05 for the common species whose cover defines plant associations. Power graphs were provided at the workshop. However, the pilot plots had

permanently marked subplots (decreasing interannual variability), a feature that is not part of FIA plots.

Power Analysis. Vegetation cover is highly variable in response to interannual variability in climate. It is also not expected to show biologically significant change (in response to changes in climate and atmosphere) except on a decadal scale. If plots are monitored at 10-year intervals, there is a danger that a wet year for the first measurement and a dry year for the second might look the same even if the fundamental trend in cover is positive. For this reason, the pilot data were analyzed as if the first period was characterized by the mean of three consecutive years, as was the second period. The test statistic was a t-test between periods using sites within plant association as replicates.

Temporal Sampling Frame. Andrea presented three possible panel designs for discussion based on the assumptions for OLYM that there are approximately 12 vegetation classes (domains) for which we want 10 plots, or a total of 120 plots. These plots would be monitored by NPS to supplement the FIA plots that are visited according to a (1-9)¹⁰ design.

Design 1: The simplest and best design for detecting trend would be a one-panel (1-0)^{all} design. This of course is not feasible. In fact, a reasonable goal is to complete 30 plots per year. The total sample for OLYM in this scenario would be 123 (FIA) plus 120 (NPS) equals 243 plots.

Design 2: Visit the complete sample of 10 plots in 3 vegetation classes for three years in a row, then move to 3 new classes for the next three years, etc. (in addition to the FIA plots). The total sample in OLYM would be 123 (FIA) plus 120 (NPS) equals 243 plots. This design would correspond to the trend analysis scheme using a paired mean comparison presented above, with each panel representing a vegetation class.

Panel = Veg Class	Year												
	1	2	3	4	5	6	7	8	9	10	11	12	13 ...
1	10	10	10										10
2	10	10	10										10
3	10	10	10										10
4				10	10	10							
5				10	10	10							
6				10	10	10							
7							10	10	10				
8							10	10	10				
9							10	10	10				
10										10	10	10	
11										10	10	10	
12										10	10	10	

Total: 30 plots per year

Design 3: Supplement FIA with some (1-4) panels and some (2-8) panels. Andrea's version was eventually modified by the end of the session to represent the recommended design.

	Year										
Panel	1	2	3	4	5	6	7	8	9	10	11 ...
1	20					20					20
2		20					20				
3			20					20			
4				20					20		
5					20					20	
1	5	5									5
2		5	5								
3			5	5							
4				5	5						
5					5	5					
6						5	5				
7							5	5			
8								5	5		
9									5	5	
10	5									5	5

Total: 30 plots per year

Discussion Points

Sample Stratification

Trent McDonald revisited the more general questions that pertain to selecting an optimal survey design and the similarity between stratified random and unequal probability designs: The first question is simply whether to use an equal or unequal probability sample. If you choose unequal, then the choice is among a stratified random sample, an intensified systematic sample, or unequal probability sampling. These are essentially three tools to achieve the same objective of distributing the sample adequately to represent all vegetation classes. Regardless of tool, you can post stratify your sample into domains for reporting (e.g., vegetation class).

Although variable probability sampling was favored for bird monitoring when the alternative was strict boundaries of accessibility, the discussion for vegetation focused on drawbacks of this method. Problems include the inflexibility of selection probabilities once they are set and the sample selected (similar to the fixed nature of strata). There was also a general concern regarding the use of highly variable inclusion probabilities. For example, if a low probability site is selected, then its weighting will cause it to contribute disproportionately to the mean and skew the trend estimate if it deviates from the other plots in its domain. If the probability differences are minor, then the concern is probably small. Probability sampling is also costly because common domains that are down-weighted to accommodate more rare domains may have small sample sizes relative to their area and therefore contribute relatively little information to the overall estimate despite representing a large space. These comments were

part of a more general plea to think carefully about using unequal probability designs given the complexity they introduce.

The recommended approach to stratification for vegetation monitoring was to use any means to distribute the sample evenly among the 12 targeted vegetation classes (domains). Fundamentally, we want to assign probability to get proper representation of each vegetation class, and we need to spread the sample evenly among them using strata that will not change. Some argued that it does not matter whether you use modeled vegetation types or precipitation, elevation, etc. to assign probabilities because any of these will probably distribute the sample to include less common vegetation types. Others strongly argued that modeled vegetation types should not be used to assign probabilities. Andrea is unwilling to work with existing vegetation maps for stratification, including modeled vegetation, due to inaccuracies. Consequently it is sensible to use exposure, elevation, and precipitation while hoping for sufficient representation of domains. When push comes to shove financially, we may decide that not all vegetation classes need to be represented.

In the case of vegetation, considering access as a weighting factor when assigning probabilities was discouraged by the group even though the travel-time map shows that there are major differences in the costs to access different parts of OLYM. Reasons include the observation that access issues may change eventually (i.e., some new technology could eliminate many access issues). Also using access as a selection factor in the park design, but not in the FIA design is inconsistent. On the other hand, multiple frames can be used in the same analysis.

One recommendation was to remove access from the design, then select an over-sample (i.e. more points than it is possible to monitor) knowing that some will be inaccessible. In the future, if accessibility changes, those sites could be added back into the sample. Over-sampling by 30% (which turned out to be a good estimate of how many would be inaccessible) at Grand Canyon led to the required number of plots for analysis and the remainder were crossed off “for now” and retained on the list for later. The advantage is that plots are excluded in an ordered process. Another possibility is to leave access out, draw the sample, and not sample the few that are unsafe to access. However, if some of the chosen plots are not used or are added later, the plots should be treated as missing data until data are actually collected. Also, experience has shown that this approach leads to ad hoc unequal probability sampling.

Given that Andrea was still concerned that access would need to be addressed, another suggestion was to use FIA plots in inaccessible areas to test whether accessible and inaccessible areas differ. If they are different and NPS wants to know about the difference, funds must be provided to monitor inaccessible areas. Alternatively, if the primary objective were to evaluate individual species independent of plant association, perhaps the data would include an adequate sample size.

Comments related to panel membership (panel not equal to domain):

Biologically meaningful analysis of vegetation plots requires that data be analyzed by domains representing vegetation classes. Otherwise, changes in dry areas, for example, may cancel those in wet areas if they are changing oppositely, appearing as zero trend. In addition, it is not desirable or possible to compare across vegetation classes because all vegetation types do not

include the same species and only a few species in each vegetation class will provide sufficient data for trend detection.

If we build from the FIA sample with its interpenetrating design, it is important to resolve the issue of domain members being scattered across different panels, which scatter their representation across years. For example, one cannot compare changes between plots that are measured in years 1 and 11 with those measured in years 2 and 12, or 5 and 15. Consequently, each year's sample must include sufficient numbers of each domain to produce adequate power to detect change. If there are 12 vegetation types and 10 plots are required for power (Andrea's pilot power analysis indicated 7-9), then 120 plots would be required per year, and this is infeasible even if NPS attempts to merely supplement FIA. These problems are unique to the analysis of domains (vegetation types); for any higher-level analysis, an interpenetrating design is fine.

Several solutions were offered. One could increase the number of samples per year, effectively guaranteeing that each domain (vegetation class) is included in each panel, or reduce the number of domains (i.e., consolidate vegetation classes). Another option is to use domains to assign membership to panels. For example, all sites in domains 1 & 2 could be sampled in panels 1&2, then not sampled for some interval. The down side of this approach is that the revisit interval for the domain type would be decadal. Therefore, the ability to detect trend for all domains will take many years (i.e., a minimum of 20 years). Also, it would be very difficult to determine the correct domain for each site before visiting it. Finally, if domains are the same as panels, you need to think about what error term is appropriate for the analysis. Temporal variance (year to year variation) or the spatial x temporal interaction could be used as the error term to judge the significance of trends. If salmonberry occurred in only one panel, you would not have these error terms.

The group also discussed that inference about change in those domains not represented in every panel could be analyzed using a model-assisted approach. Design-based analyses draw randomness from site selection itself and incorporate selection probabilities; model-assisted analyses ignore selection probabilities and evaluate variance itself (ANOVA, etc.) A model might be used to estimate trend (rate of change) at each site then use design-based estimation using the site trend as the response. More complex analyses using selection probabilities are possible. Participants agreed that a design-based inference is preferable to the model-assisted analysis, but model-assisted inference is acceptable if the model assumptions are defensible.

Revisit Design:

Concern was expressed about using the FIA revisit scheme where 10% of the sample is visited every year, then repeated after 10 years. This scheme requires 20 years before all plots have been visited twice to provide trend data and it misses changes happening at the sub-decadal time scale. It is doubtful that a decadal-scale approach is adequate for providing management information. Therefore the group concluded that the design should include some revisit plan on a shorter interval to provide the opportunity for earlier trend detection. At the same time it would be good to have some component overlapping with FIA because in 100 years such a design would have tremendous power.

One recommendation was to use a five-year revisit plan because it provides more frequent revisits but stays in synchrony with the FIA schedule (i.e., a 10 year revisit plan) or to use both 5- and 10-year revisits, which would increase spatial coverage to describe status of rarer domains. This is an important compromise that preserves the spatial extent of the sample while providing better trend information and connectedness between years. Shorter return intervals also take site variability (expected to be quite large) out of the trend estimation sooner, meaning much higher power to detect trend (see Scott Urquhart's talk).

Without some repeat visits in consecutive years, the design is not connected and therefore cannot account for interannual variation. Interannual variation is expected to be high because vegetation cover is very sensitive to annual weather conditions. It may not be necessary to revisit all sites for consecutive years, and any one site for more than two consecutive years. As long as some sites are measured in two consecutive years, you can estimate interannual variability, but this assumes annual correlation among sites and a model-assisted approach. Making few consecutive revisits is desirable for vegetation because permanent plots are subject to a large response burden.

Analysis:

Andrea described data analyses where plots were visited for three years to describe time period 1, followed by three more years after an interval of 8 years. Trend was estimated as a period means comparison between the two time periods. The variance among trends for all sites within a domain was used to test whether the mean trend for a domain was significant. This generated a discussion about how variability influences power of trend estimates. The existing analysis treats each site independently. This idea removes the site effect, but cannot be conducted until a revisit sequence is complete. The site-specific trends are then averaged to say something about the domain. Unequal probabilities could be applied to each individual trend before aggregating the values. This approach makes sense because the variance between sites is not informative, and it is expected to be quite high (e.g., 2% cover versus 80% cover of the same species for plots in the same domain). Some disagreed with this approach, saying that it is better to use a linear regression metric instead of a period comparison because averaging three years (period means comparison) is less powerful than linear regression. This contradicts the actual power analysis of Andrea's data, based on simulations, which showed virtually no difference in power between the two methods.

Recommended Approach:

Spatial Design. Despite the caution against using variable probabilities, the group recommended that plots should be allocated using variable probabilities among physical strata that approximate vegetation classes. Using access as a stratum was discouraged. Incorporate FIA plots as much as possible.

Temporal Design. The recommended design was a $[(1-4)^5(2-8)^{10}]$ panel design that includes 20 plots in the first 5 panels and 5 plots in the second 10 panels. A total of 30 plots would be visited per year. This design strikes a balance among maximizing spatial distribution, coordinating with the FIA schedule, providing a shorter revisit interval, and creating an interconnected design. (See Design 3 chart above in this section).

One tool for deciding on a specific temporal design is to consider what you want to estimate, then run the analysis with artificial data. You can also compare the projected precisions under alternate designs. For example, consider Andrea's period-means approach with only two years per period and a reduced revisit interval (so change estimates will be available sooner). The following example is based on artificial data with 1 site per panel generated by adding the site and year indices and an error value (a random number <1).

	Year													Trend (annual rate)
Site	1	2	3	4	5	6	7	8	9	10	11	12	13	1.033
1	2.7	3.5					8.9	9.7						1.025
2		4.6	5.1					10.6	11.4					1.025
3			6.9	7.3					12.4	13.0				0.933
4				8.0	9.7					14.1	15.7			1.008
5					10.8	11.0					16.9	17.0		1.008
6						12.4	13.6					18.4	19.5	0.992

The trend equals: [(mean of the last 2 years – mean of first 2 years)/6 years] to estimate an annual rate. The mean trend is 1.000 stems per year with standard error 0.015 and 95% confidence interval (0.962, 1.038). The design is connected, although this property is not used in the example. Because it is connected, you can estimate the difference between years 1 and 5 as $(2.7-3.5+4.6-5.1+6.9-7.3+8.0-9.7) = 3.4$.

Membership Design. There was no obvious solution to the problem of the membership design. Possible approaches include making panels equal to vegetation classes, but that will not work well with the FIA design and would not be interpenetrating. Vegetation classes could be combined into larger categories that are likely to be well represented in the FIA sample each year and supplement those that will not be covered by FIA. An interpenetrating design will make the period mean comparison unworkable, and although model-assisted analyses were discouraged, they should possibly be investigated.

Developing Stream Habitat Monitoring for NCCN (presented by Reed Glesne, North Cascades National Park Complex)

Background

At an earlier stage of the NPS monitoring program, prior to the development of networks, NOCA was chosen to develop protocols for monitoring streams and lakes. Reed Glesne has spent many years developing the monitoring program for NOCA, especially collecting data to develop benthic macroinvertebrate (BMI) models for the park. The example projects he provided for this workshop include stream fish and stream water quality, both chemical and biological, in wadeable streams.

It was emphasized during the workshop that the approach suggested for NOCA was not appropriate for the other parks in the NCCN. Other parks have different levels of precipitation and geomorphology, which affect the hydrologic system. Also, low gradient wadeable streams are not abundant in OLYM and MORA and the appropriate number of regions for subsampling differs from park to park. It was recognized that more work is needed to expand Reed's approach to the other parks.

The monitoring objectives for summer stream resident fish are:

- Determine status and trends of **total abundance** and **frequency of occurrence** of target fish species from pool habitats for selected sample reaches in four major regions of NOCA, and make park-wide extrapolations. Target species include bull-, rainbow- and cutthroat trout, and young-of-the-year Coho.

Both measurements would be accomplished by sampling stream reaches. Total abundance would be estimated using a combination of first stage single pass and/or multi-pass diver surveys calibrated by electrofishing as population estimates from a systematic sample with a random systematic start of pools from each reach. Additional habitat attribute data would be collected during surveys. Frequency of occurrence would be monitored using single pass diver surveys of a random start systematic sample of 100 m stream segments in each reach. Sampling would be limited to stream reaches with gradients 0-4%. There are approximately 150 of these reaches in NOCA, constituting approximately 50% of stream habitat, and they vary from 0.5 to 1.5 miles in length. Low gradient reaches were chosen because they represent the major portion of fish habitat, they are areas where responses to change are expected, and they are accessible.

Reaches would be selected by a random start systematic sample within a different park region each year, where the regions are similar in size. Specifically, sample reaches are selected starting at a randomly chosen reach at the downstream end of one of the stream systems within a particular park region. Reaches are selected systematically working upstream within that system and then to the lowest downstream reach of the next adjacent stream system, and so on. For park-wide extrapolation, a sample of randomly picked reaches from the park-wide collection of sample reaches is selected and surveyed annually. It may be necessary to intensify sampling in some stream systems that are of particular management concern. A sample of 25 to 30 reaches would be sampled per year.

The objective for stream water quality and biomonitoring is:

- Determine status and trends of water quality and biological condition of wadeable streams as indicated by BMI metrics and selected physical and chemical parameters.

Monitoring would be accomplished by sampling 4 to 5 riffles for BMI in the lowest accessible downstream reach in randomly selected 7th field HUCs (i.e., watersheds with area of 2000 – 10,000 acres). BMI samples are pooled for analysis and are representative of the reach. Both multi-metric and predictive models are used to evaluate BMI data. Other physical and chemical attribute data will be collected in the sample reach, including explanatory attributes associated with BMI data, site level disturbances, habitat conditions (e.g., wood, canopy cover, channel

dimensions) and the mandatory NPS suite of basic chemical parameters. Some sites would be sampled at 3-5 year intervals and others annually. Flexibility to intensify sampling within regions of parks is needed.

The number of 6th and 7th field HUCs for the large parks in the network are:

Park Unit	6 th Field HUCs (10k-40k acres) ¹	7 th Field Hucs (2k-10k acres) ¹
MORA ²	12	50
NOCA ²	50	130
OLYM	71	142

¹all that touch the park borders

²currently in draft form.

Reed noted that his methods for sampling BMI are compatible with those of AREMP and EMAP, other monitoring programs in the region, but the sampling designs are different. The importance of the level of compatibility among these three protocols became the primary discussion topic for aquatic monitoring.

Reed was asked why the panel is restricted to one region when each panel (annual sample) could cover all regions to provide and estimate for the entire park every year. As it stands the sample provides estimates for each region for each year. Reed answered that the Park is interested in estimates for each region independently, and the logistics of a regional sample are simpler. Therefore, a block design is preferable in this case.

Discussion Points

The discussion of aquatic sampling never addressed the temporal sampling frame. Instead it focused on objectives, target population, response design, and sample unit. Generally Reed was challenged to justify his approach against that of EMAP.

Objectives/Approach

The group recognized similarities between Reed's monitoring indicators and those of EMAP and wondered why EMAP protocols and sampling scheme are not adopted at least as a starting point. Reed stated that while the NPS program will definitely connect biological, chemical and habitat measurements like EMAP, it differs in some objectives, as well as target population and sample units:

- Estimates of abundance for fish species are not included in EMAP.
- NPS limits the target population to 0-4% gradient streams for cost reasons.
- Monitoring by blocks (region) is also a cost-saving measure. (Regions instead of annual park-wide assessment)
- NPS wants stream-level information for fish, not by HUCs.
- NPS BMI and chemistry are done at the 7th field HUC, rather than 6th field HUC level.
- NPS wants quantitative assessment of streams rather than categories of condition
- NPS wants to monitor stream reaches instead of points. NPS needs to detect much smaller changes than EMAP because NPS monitoring is focused on unmanaged, fairly pristine areas. To detect a small change, one must take a lot of samples or have a very sensitive indicator.

The group emphasized the need for detailed objectives (ala Tony Olsen's examples) before undertaking any further fieldwork or protocol development. The objective of determining status and trends of frequency and abundance of fish species or water quality do not identify what will be estimated. For example, one could estimate trend (average rate of change) by species, or annual means by species, among other things. They recommended that he consider using some aspects of the EMAP program because it also includes chemical and benthic monitoring. They felt that some efficiency could be achieved with collaboration.

Target Population

EMAP defines the target population as the entire stream network, which is a plan that will likely be followed by the states. The NPS approach is to stratify the park by gradient and limit the sample (and inference) to low gradient areas. Model-based park-wide inference could be made using the assumption that low gradient reaches represent the entire park. An advantage may be that you can collect more data in the relatively level areas than you could if you try to extend the sample over the entire river network. However, by considering low gradient areas to be 'response' zones, the implication is that the sample will over-estimate change for the whole park. Selecting the lowest accessible downstream reaches will bias the results if there is an upstream gradient and you will be unable to say anything about large areas of the park. Reed was strongly encouraged to increase his target population to at least 0-8% gradient streams. While it is not always possible to sample some of the steepest gradients because snorkeling isn't feasible, collecting the benthic and chemical data is possible. The justification for the 4% limit is that NPS can't afford to sample the entire stream network, and areas with 0-4% gradient have the highest density of fish species, but the validity of this reason was questioned. Experts felt that one could incorporate cost into the design and have inference to more of the park for the same cost. Following this discussion, Reed agreed to consider stream reaches up to 8% gradient, maybe also including riffles in addition to pools.

Response Design

The group noted that EMAP looks at all species of fish, while NPS looks at only four (rainbow-, bull-, and cutthroat trout, and young Coho). Reed pointed out that this isn't a problem because he is looking at only four species for abundance, but will get relative abundance for all species. Population totals cannot be generated for species other than the four because they are too rare.

It was also noted that EMAP has two protocols, one for responses in wadeable streams, and another one for non-wadeable ones. The non-wadeable protocols may be a way to address the unmet needs of other parks, especially OLYM.

Sample Unit: Point versus Reach

Reed's sample unit is a stream reach, whose length is defined as 40 times the average width of the stream. Because of the dependence of the definition on stream width, sample units do not have uniform sizes and range from 0.5 to 1.5 miles long. Pools are chosen for sampling using a random start systematic sample of pools within a reach (e.g., every third pool is chosen). Pools are not the target population; instead, pools are used to make inference to the reach. The average number of pools per mile is 20, but it can be as high as 40. The design essentially boils

down to a two-stage approach – first the reach is selected, then pools are randomly selected within the reach.

Generous discussion was given to problems associated with different segment lengths; the relationship between stream width, stream morphology and the subsequent definition of stream reach; and the relationship between segment length and pool density. The problems focused around defining the ultimate unit of measurement and how measurement units could be standardized. Without equal sized units, inference is questionable and there may be bias if the stream lengths vary greatly. Even if the sample is not biased, it is likely inefficient. Reed countered that all the streams come into the sample with the same probability, and the total number of fish is estimated without bias if it is the total per stream (not per mile). All the pools are assigned to a reach uniquely and the breaks for reaches come from maps and aerial photos. They are expected to remain stable for at least the next 10-20 years.

Several alternatives to unequal reach lengths were explored:

1) Sample pools, instead of reaches. This idea was discarded because pool locations change as river morphology changes due to floods. Then it was noted that if pools can change position, it is possible for the ones on the border to change reaches over time. Reed said there are simple rules to deal with boundary pools. Rules to accommodate special circumstances will be required regardless of design (e.g., encountering a waterfall, change in gradient, etc.) Inclusion probabilities of pools change: isolated pools have a lot more probability of being included.

2) Make all reaches the same length. Reed said that 30% of the reaches are one-half mile in length and the others are longer. If all of the reaches were one-half mile, many would not have enough pools. Moreover longer streams are geomorphologically different than small ones, and thus provide different types of habitat. Reed pointed out that no correlation between fish number and stream length has been observed, maybe because length of stream might not be related to pool habitat (i.e., short streams may have a lot of pools). Also, unequal stream length could be taken into account by rescaling by stream length.

3) Perhaps the relative contribution of each stream reach to the overall estimates could be normalized by standardizing the number of pools per reach.

EMAP uses point sampling, where points are systematically located, and all values, fish as well as water quality and macroinvertebrates, are collected at the same time. Fish abundance is determined in fixed distance on either side of the point. This contrasts with Reed's approach, which defines stream reaches in two different ways, one for fish and one for water quality and macroinvertebrates. The EMAP sampling scheme eliminates the need for an unbiased estimator of macroinvertebrate density and water quality that accounts for different reach lengths (another sampling complexity). With points, samples of macroinvertebrate and chemical parameters taken at the point are unbiased. However, it was pointed out that bias cannot be avoided with either reach-based or point-based sampling. The bias associated with using reaches comes from the variability in reach length. The bias of points is associated with how points that fall near the end of the reference frame are handled. For example, if the protocol is to go 50 m up and down the stream from the point, and the point falls at the end of a

sample frame, you may choose to sample 100 m upstream. The result is that the ends of the reaches have a slightly higher probability for sampling. However, this could be solved by weighting the data from boundary pools by the fraction of each that is in the reach. For example, if 60% of the pool is in the sample frame, the observation for that pool would be 60% of the total pool count. Another source of error is pools on the margins moving in and out of the sample frame, but this could be resolved by some simple rules.

There was much discussion about whether Reed could/should use the point-based approach, and how he would select the area around the point – would it be based on the number of pools or on a fixed distance? Most participants advocated using a point-based approach for water quality and macroinvertebrates with fish sampled in all pools within a fixed distance around the point. Finding no fish in the sampled pools would be a legitimate result.

Recommended Approach

It was argued that the logical starting point for stream monitoring is to use EMAP or AREMP (associated with Northwest Forest Plan monitoring) protocols and standard monitoring designs. These could be modified to include estimates of abundance if required. It was noted, however, that EMAP is a program to develop monitoring tools, not a long-term monitoring program. Consequently, NPS should look towards the states and region to standardize the methodology, expecting that these programs are likely to be based on EMAP.

Conclusions

One of the challenging requirements of NPS long-term ecological monitoring is to simultaneously describe status and trends. The statistical community has introduced split panel temporal designs as a means to meet this challenge. This workshop explored the properties, limitations, and trade-off among possible split panel and other temporal sampling designs. A few key points are:

- Estimates of status increase with the total number of sites sampled.
- Estimates of trend increase with decreasing measurement interval for each site.
- For equal effort, estimating status comes at a price for estimating trend and vice versa.
- Statistically connected designs are especially costly for describing status, but they allow comparison between individual years when annual patterns of change are important to understand.
- Temporal sampling frames and spatial sampling frames are closely intertwined because decisions about panel membership are influenced by the spatial distribution of sites.
- Spatial distribution of sampling sites will often be based on unequal probability sampling to distribute sites cost effectively, safely, or efficiently among common and rare populations, and the possible methods come with their own set of trade-offs.

A number of alternative statistical survey designs were discussed. To choose among them for any project, it is necessary to identify very specific objectives (i.e., parameters to be estimated) and the analyses to be used. This is the only way to compare the relative advantages of the alternative designs. For example, if you are only interested in decadal changes, there is no need to measure plots in successive years to estimate annual changes; if status is more important than trend you can lengthen revisit intervals thereby increasing the number of sites visited in a

given amount of time. An important step is to carefully examine each potential design to see what can be estimated precisely, what can be estimated imprecisely, and what cannot be estimated with the analyses you plan to use. There is a danger that if you try to estimate too many different parameters you will not do any of them well. The importance of having specific objectives before proposing and evaluating the temporal and spatial components of survey designs cannot be overemphasized.

Acknowledgements

I am grateful to all of the participants of this workshop for their unflagging attention and stimulating discussion. Special acknowledgment is due to the presenters, Reed Glesne, Kurt Jenkins, Tony Olsen, Trent McDonald and Scott Urquhart. This summary would not have been possible without the notes taken by Katherine Bierne, Mike DeBacker, John Boetsch, Carrie Donnellan, Jerry Freilich, Cat Hoffman, Giovanna Ranalli, Barbara Samora, and Alice Shelly, and their help is greatly appreciated. Kurt Jenkins, Paul Geissler and Trent McDonald made helpful comments on earlier drafts of the summary.

References

Buckland, S. T., D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers and L. Thomas. 2001. Introduction to distance sampling: estimating abundance of biological populations. Oxford University Press. Oxford, UK. 432 pp.

Elzinga, C. L., D. W. Salzer, J. W. Willoughby. 1998. Measuring and monitoring plant populations. BLM Technical Reference 1730-1.

Jenkins, K. J, A. Woodward, and E. G. Schreiner. 2003. A framework for ecological monitoring in Olympic National Park: prototype for the coniferous forest biome. US Geological Survey Information and Technology Report ITR 2003-006.

McDonald, T. L. 2003. Review of environmental monitoring methods: survey designs. Environmental Monitoring and Assessment 85:277-292.

Stevens, D. L., Jr. 1997. Variable density grid-based sampling designs for continuous spatial populations. Environmetrics 8: 167-195.

Urquhart, N. S., S. G. Paulsen and D. P. Larsen. 1998. Monitoring for policy relevant regional trends over time. Ecological Applications 8: 246-257.

Woodward, A. and K. J. Jenkins. 2001. Summary of a statistics workshop – April 2-3, 2001. <http://fresc.usgs.gov/olympic/workshops.html>.

Appendix A. Participants of the Temporal Sample Frame Workshop – 12-14 Nov 2003

Steve Acker Olympic National Park 600 East Park Avenue Port Angeles, WA 98362 360-565-3073 Steve_Acker@nps.gov	Steve Fancy National Park Service-WASO 1201 Oak Ridge Dr. #200 Ft. Collins, CO 80525 970-225-3571 Steven_Fancy@nps.gov	Patti Happe Olympic National Park 600 East Park Avenue Port Angeles, WA 98362 360-565-3065 Patti_Happe@nps.gov
Mignonne Bivin North Cascades Nat. Park Complex 7280 Ranger Station Rd Marblemount, WA 98267-9702 360-873-4590 x. 58 Mignonne_Bivin@nps.gov	Steve Fradkin Olympic National Park 600 East Park Avenue Port Angeles, WA 98362 360-928-9612 Steven_Fradkin@nps.gov	Cat Hoffman Olympic National Park 600 East Park Avenue Port Angeles, WA 98362 360-565-3060 Cat_Hoffman@nps.gov
John Boetsch Olympic National Park 600 East Park Avenue Port Angeles, WA 98362 360-565-3064 John_Boetsch@nps.gov	Jerry Freilich Olympic National Park 600 East Park Avenue Port Angeles, WA 98362 360-565-3082 Jerry_Freilich@nps.gov	Bob Hoffman USGS-FRESC 777 NW 9 th Street, Suite 400 Corvallis. OR 97330 541-750-1013 Robert_Hoffman@usgs.gov
Sam Brenkman Olympic National Park 600 East Park Avenue Port Angeles, WA 98362 360-565-3081 Sam_Brenkman@nps.gov	Paul Geissler USGS – BRD C/o BARC – East Bld. 308 Rm. 126 10300 Baltimore Avenue Beltsville, MD 20705 301-497-5780 Paul_Geissler@usgs.gov	Roger Hoffman Olympic National Park 600 East Park Avenue Port Angeles, WA 98362 360-565-3062 Roger_Hoffman@nps.gov
Tamre Cardoso Terrastat Consulting Group 10636 Sand Point Way NE Seattle, WA 98125 206-527-9857 tamre@blarg.net	Scott Gende Glacier Bay Field Station 3100 National Park Rd. Juneau, AK 99801 907-364-1591 Scott_Gende@nps.gov	Gail Irvine USGS- Alaska Science Center 1011 E. Tudor Rd. Anchorage, AK 99503 907-786-3653 Gail_Irvine@usgs.gov
Jean-Yves (Pip) Courbois University of Washington Box 354322 Seattle, WA 98115-4322 206-616-9439 pip@stat.washington.edu	Reed Glesne North Cascades Nat. Park Complex 810 State Route 20 Sedro Woolley, WA 98284- 9314 360-856-5700 x. 369 Reed_Glesne@nps.gov	Kurt Jenkins USGS-FRESC-Olympic Field Statn. 600 East Park Avenue Port Angeles, WA 98362 360-565-3043 Kurt_Jenkins@usgs.gov
Mike DeBacker Wilson's Creek National Battlefield 6424 West Farm Rd. 182 Republic, MO 65738-9514 417-732-6438 x. 269 Mike_DeBacker@nps.gov	Shelley Hall Olympic National Park 600 East Park Avenue Port Angeles, WA 98362 360-565-3066 Shelley_Hall@nps.gov	Bob Kuntz North Cascades Nat. Park Complex 810 State Route 20 Sedro Woolley, WA 98284- 9314 360-856-5700 x. 368 Robert_Kuntz@nps.gov

Phil Larsen US Environmental Protection Agency 200 SW 35 th Street Corvallis, OR 97333 541-754-4362 larsen.phil@epa.gov	Jim Schaberl Mount Rainier National Park Tahoma Woods Star Route Ashford, WA 98304-9751 360-569-2211 x. 3373 Jim_Schaberl@nps.gov	Andrea Woodward USGS – FRESC C/o WFRC 6505 NE 65 th Street Seattle, WA 98115 Andrea_Woodward@usgs.gov
Trent McDonald Western EcoSystem Tech., Inc 2003 Central Avenue Cheyenne, WY 82001 307-634-1756 tmcDonald@west-inc.com	Alice Shelly Terrastat Consulting Group 10636 Sand Point Way NE Seattle, WA 98125 206-362-3299 alice@blarg.net	
Karen Oakley USGS-BRD 1110 East Tudor Rd. Anchorage, AK 99503 907-786-3579 Karen_Oakley@usgs.gov	Rodney Siegel Institute for Bird Populations PO Box 1346 Point Reyes, CA 94956 415-663-2051 rsiegel@birdpop.org	
Tony Olsen US Environmental Protection Agency 200 SW 35 th Street Corvallis, OR 97333 541-754-4790 olsen.tony@epa.gov	Don Stevens Oregon State University 44 Kidder Hall Corvallis, OR 98 541-737-3587 stevens@stat.orst.edu	
M. Giovanna Ranalli Department of Statistics Colorado State University Ft. Collins, CO 80523-1877 970-491- ranalli@stat.colostate.edu	N. Scott Urquhart Department of Statistics Colorado State University Ft. Collins, CO 80523-1877 970-491-3834 nsu@stat.colostate.edu	
Regina Rochefort North Cascades Nat. Park Complex 810 State Route 20 Sedro Woolley, WA 98284-9314 360-856-5700 x. 254 Regina_Rochefort@nps.gov	Samantha Weber Mount Rainier National Park Tahoma Woods Star Rte. Ashford, WA 98304-9751 360-569-2211 x. 3395 Samantha_Weber@nps.gov	
Barbara Samora Mount Rainier National Park Tahoma Woods Star Route Ashford, WA 98304-9751 360-569-2211 x. 3372 Barbara_Samora@nps.gov	Bob Wilkerson Institute for Bird Populations PO Box 1346 Point Reyes, CA 94956 415-663-1436 bwilkerson@birdpop.org	

Appendix B. Additional Examples Comparing Designs from Scott Urquhart

TEMPORAL DESIGN 1 ALWAYS REVISIT

	TIME PERIOD (ex: YEARS)												
PANEL	1	2	3	4	5	6	7	8	9	10	11	12	13 ...
1	X	X	X	X	X	X	X	X	X	X	X	X	X

TEMPORAL DESIGN 2: NEVER REVISIT

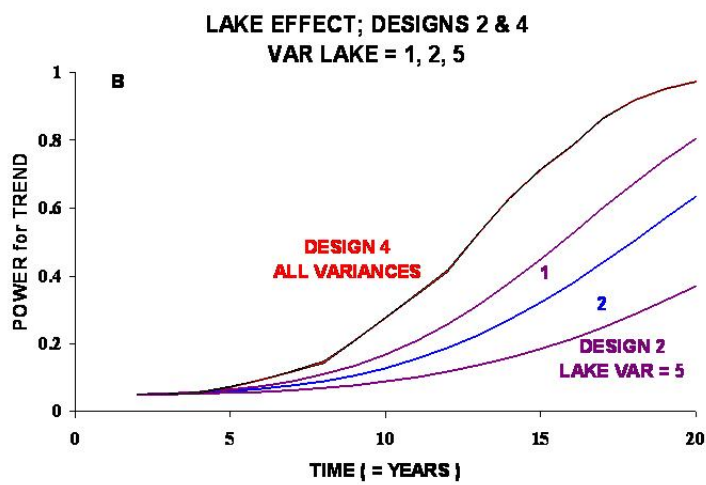
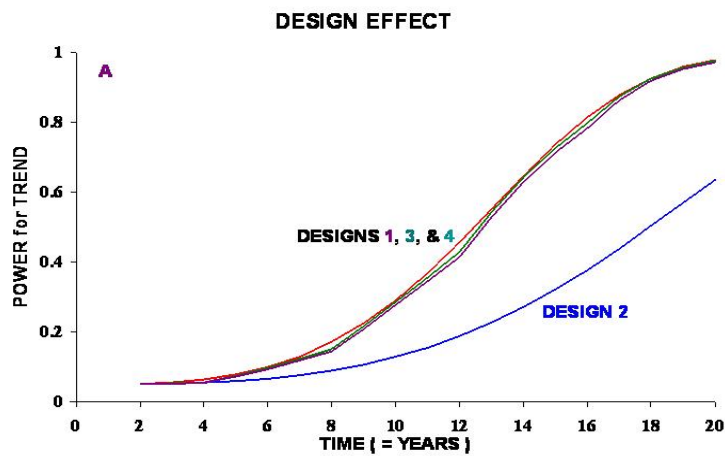
	TIME PERIOD (ex: YEARS)												
PANEL	1	2	3	4	5	6	7	8	9	10	11	12	13 ...
1	X												
2		X											
3			X										
4				X									
5					X								
6						X							
7							X						
8								X					
9									X				

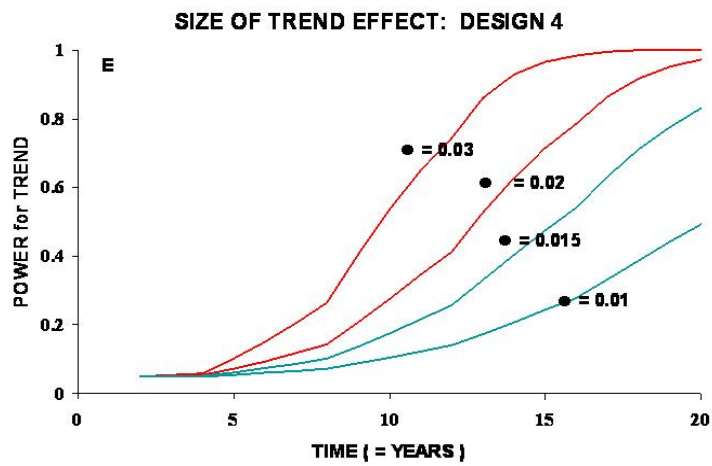
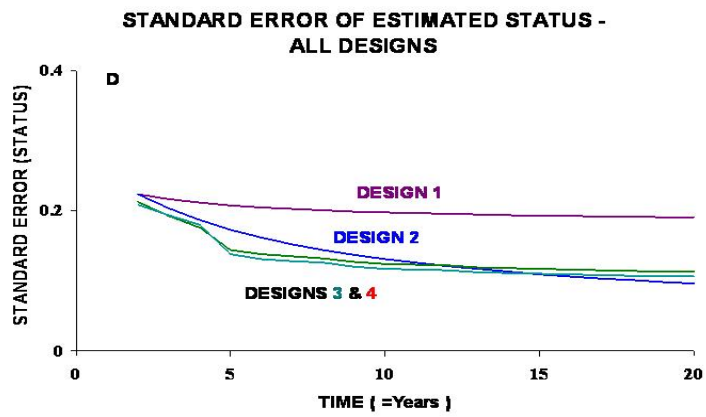
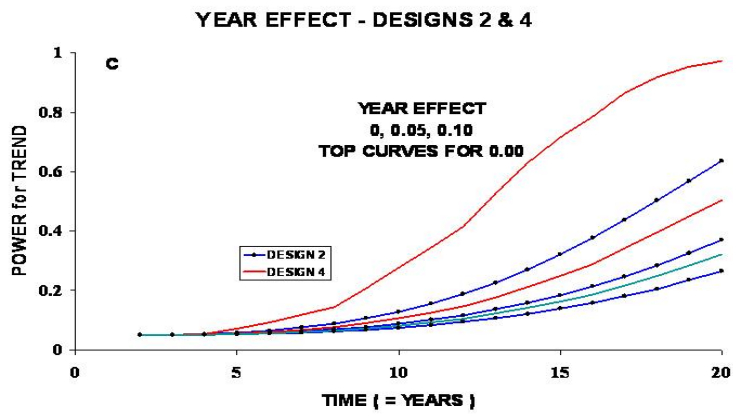
TEMPORAL DESIGN 3: AUGMENTED SERIALY ALTERNATING

	TIME PERIOD (ex: YEARS)												
PANEL	1	2	3	4	5	6	7	8	9	10	11	12	13 ...
0	X	X	X	X	X	X	X	X	X	X	X	X	X
1	X				X				X				X
2		X				X				X			
3			X				X				X		
4				X				X				X	

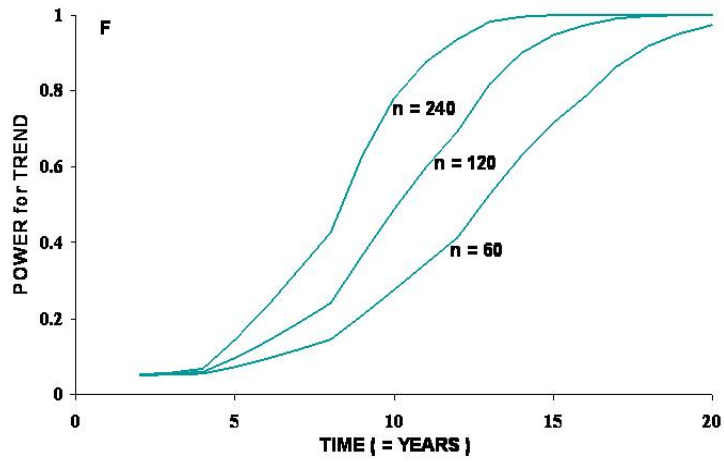
TEMPORAL DESIGN 4: SPLIT PANEL
SERIALLY ALTERNATING
 PLUS SERIALY ALTERNATING WITH CONSECUTIVE
 YEAR REVISITS

PANEL	TIME PERIOD (ex: YEARS)													...
	1	2	3	4	5	6	7	8	9	10	11	12	13	
1	X				X				X				X	
1A	X	X			X	X			X	X			X	
2		X				X				X				
2A		X	X			X	X			X	X			
3			X				X				X			
3A			X	X			X	X			X	X		
4				X				X				X		
4A				X	X			X	X			X	X	

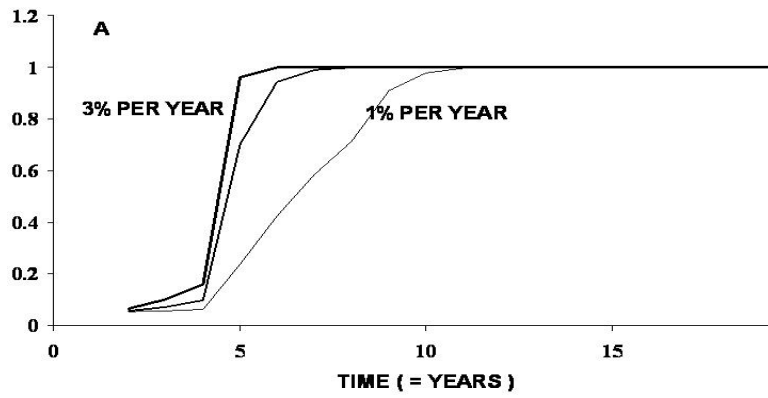




SAMPLE SIZE EFFECT - DESIGN 4



SECCHI DEPTH



ln (CHLOROPHYLL a)

